

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2002, 12:37:23 ; Search time 23.8 Seconds  
(without alignments)  
497.972 Million cell updates/sec

Title: US-09-731-816-4  
Perfect score: 864  
Sequence: 1 NSARAVLSAFHHTLQLGP.....SIDKQAKLLLPNDPAGP 160

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues  
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	864	100.0	160	21 AAY53891	Partial amino acid
2	835	96.6	173	21 AAY53893	Partial amino acid
3	835	96.6	185	21 AAY70656	Mature human trans
4	835	96.6	186	21 AAY70655	Mature human trans
5	835	96.6	187	21 AAY70654	Mature human trans
6	835	96.6	187	21 AAY70663	Mature human trans
7	835	96.6	202	21 AAB07595	A human interleukin
8	835	96.6	202	21 AAB07689	Human transforming
9	835	96.6	202	21 AAY70653	Human transforming
10	835	96.6	209	21 AAY70662	Human transforming
11	831	96.2	202	22 AAU04953	Human Interleukin

12	719	83.2	151	21	AAB07594	A human interleukin
13	719	83.2	151	21	AAB07690	A rat interleukin
14	707	81.8	183	21	AAY70658	Mature murine tran
15	707	81.8	205	21	AAB07597	A murine interleuk
16	707	81.8	205	21	AAB07691	A murine interleuk
17	707	81.8	205	21	AAY70657	Murine transforming
18	707	81.8	205	22	AAU05341	Mouse zcyto7 prote
19	460	53.2	103	21	AAB07688	A human interleukin
20	303	35.1	57	21	AAY70666	Human transforming
21	298	34.5	54	21	AAY70664	Human transforming
22	208	24.1	44	21	AAB07596	A rat interleukin
23	197.5	22.9	425	21	AAY44462	Human interleukin
24	192	22.2	197	21	AAE18911	A novel polypeptid
25	192	22.2	197	21	AAB07602	A human interleukin
26	192	22.2	197	21	AAB07684	A human interleukin
27	192	22.2	197	21	AAAY92238	Human interleukin
28	192	22.2	197	21	AAY44460	Human interleukin
29	192	22.2	197	21	AAAY53892	Amino acid sequenc
30	192	22.2	197	22	AAU04951	Human interleukin
31	192	22.2	206	21	AAAY44485	Human interleukin
32	180	20.8	34	21	AAAY70661	Human transforming
33	177	20.5	158	20	AAW85646	Murine zcyto7 matu
34	177	20.5	160	20	AAW85645	Murine zcyto7 matu
35	177	20.5	180	20	AAW28236	Mouse interleukin
36	177	20.5	180	20	AAW22198	Murine EDIRF 1 pro
37	177	20.5	180	20	AAW85617	Murine zcyto7. Mu
38	177	20.5	180	21	AAB07593	A murine interleuk
39	177	20.5	180	21	AAB07687	A murine interleuk
40	176.5	20.4	408	21	AAAY44461	Human interleukin
41	174	20.1	128	20	AAW85643	Human zcyto7 matur
42	174	20.1	130	21	AAAY93975	Amino acid sequenc
43	174	20.1	151	20	AAW85641	Human zcyto7 matur
44	174	20.1	153	20	AAW85642	Human zcyto7 matur
45	174	20.1	154	20	AAW85640	Human zcyto7 matur

## ALIGNMENTS

RESULT 1

AAV53891

ID AAY53891 standard; Protein; 160 AA.

XX AAY53891;

XX AAY53891;

XX 13-MAR-2000 (first entry)

XX Partial amino acid sequence of human interleukin-22.

XX Human; Interleukin-22; IL-22; IL-21; Immune system disorder;

XX Immune cell chemotaxis; haematopoietic cell disorder;

XX haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;

XX respiratory problem; organ rejection; graft-versus-host disease; GVHD;

XX Inflammation; hyperproliferative disorder; tissue regeneration;

XX embryonic stem cell differentiation; embryonic stem cell proliferation;

XX haematopoietic lineage; allergic asthma.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 26

XX Domain /note= "potential N-linked glycosylation site"

XX Domain /note= "conserved domain I"

XX Domain /note= "conserved domain II"

XX Domain /note= "conserved domain III"

XX Domain /note= "conserved domain IV"

XX Modified-site 139

XX /note= "potential N-linked glycosylation site"



	Matches	154;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	5	ARVLSAFHHTLQIGPREQARNASCPAGGRPADRRFRPPTNLRSVSPWAYRISYDPA	RP	64						
Db	30	aagvlsafhtlqlgpreqarnascpaggrpadrrfrpptnlrsvspwayrisydparyp	89							
Qy	65	RYLPEAYCLRCGCLTGLFGEEDVFRSAPVYMTPTVWLRRTPACAGGRSVYTEAYVTIPVG	124							
Db	90	rylpeayclrcgcltglfgeedvfrsapyymptvvlrrtpacaggrsvyteayvtipvg	149							
Qy	125	CTCVPPEKQADSTNSSIDKQAKLLGPNDA	PAGP	160						
Db	150	ctcvepekdaadsinssidkqaklllgpndapagp	185							

RESULT	4
AAAY70655	
ID	AAAY70655 standard; Protein; 186 AA.
XX	
XX	AAAY70655;
XX	
XX	18-JUL-2000 (first entry)
XX	
DE	Mature human transforming growth factor beta-9, Ztgfbeta-9 protein-2.
XX	
XX	Human transforming growth factor beta-9; Ztgf beta-9;
KW	Alzheimer's disease; neurodegenerative disease; Huntington's disease;
KW	amyotrophic lateral sclerosis; ALS; Parkinson's disease;
KW	peripheral neuropathy; demyelinating disease; multiple sclerosis;
KW	antiviral; cytostatic.
XX	
XX	Homio sapiens.
OS	
XX	W0200015798-A2.
PN	
XX	23-MAR-2000.
PD	
XX	17-SEP-1999; 99WO-US21677.
XX	
PF	17-SEP-1998; 98US-0154817.
XX	
PR	(ZYMO ) ZYMOGENETICS INC.
XX	
PA	Presnell SR, Taft DW, Foley KP;
XX	
PI	WPI: 2000-271436/23.
XX	
DR	N-PSDB; AAZ52195.
DR	
XX	Polynucleotides encoding a novel transforming growth factor beta-9
PT	polypeptide, designated ztgf beta-9, useful as an antiviral and
PT	antiproliferative agent
XX	
XX	Claim 6; Page 84; 97pp; English.
PS	
XX	
CC	The present sequence is the mature human transforming growth factor
CC	beta-9, designated ztgf beta-9. This is a mature sequence excluding the
CC	signal sequence extending from amino acid 17 to and including amino acid
CC	202 of ztgf beta-9. Human ztgf beta-9 was isolated from an arrayed
CC	pituitary gland cDNA plasmid library by PCR screening. This can be used
CC	to treat a variety of neurodegenerative diseases such as amyotrophic
CC	lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,
CC	Parkinson's disease and peripheral neuropathies, or demyelinating
CC	diseases including multiple sclerosis. Ztgf beta-9 peptides have
CC	antiviral activity and may also be used to regulate the proliferation,
CC	differentiation and apoptosis of neurons, glial cells, lymphocytes,
CC	hematopoietic cells and stromal cells.
XX	
SQ	Sequence 186 AA;
XX	

```
Query Match          96.6%; Score 835; DB 21; Length 186;
Best Local Similarity 98.7%; Pred. No. 1.9e-85;
Matches 154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```



QY 65 RYLPEAYCLRGCLTGLFGEEDVFRSAPYMTVVLRRTPACAGRSVYTEAYVTIPVG 124  
|||||  
Db 92 RYLPEAYCLRGCLTGLFGEEDVFRSAPYMTVVLRRTPACAGRSVYTEAYVTIPVG 151  
|||||  
QY 125 CTCVPEPEKADSNSSIDKQAKLLGPNDA PGP 160  
|||||  
Db 152 CTCVPEPEKADSNSSIDKQAKLLGPNDA PGP 187  
|||||

RESULT 7  
AAB07595  
ID AAB07595 standard; Protein; 202 AA.  
XX AC AAB07595;  
XX DT 07-NOV-2000 (first entry)  
XX DE A human interleukin (IL) 173 polypeptide.  
XX KW Interleukin; IL-17; CTLA-8; IL-170; IL-172; IL-173; IL-174; IL-176;  
XX IL-177; IL-171; cell proliferation; cancer.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Peptide 1..17  
FT Protein /note= "signal peptide"  
FT Protein 18..202  
FT Modified-site /note= "mature protein"  
FT Modified-site 50  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 50..53  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 53  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 57..61  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 57..61  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 66..69  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 66..69  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 72..75  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 72..75  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 72  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 75  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 80  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 82  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 82..84  
FT Modified-site /note= "calcium phosphorylation site"  
FT Modified-site 113  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 113..116  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 113..116  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 116  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 164..166  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 164..166  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 166..168  
FT Modified-site /note= "calcium phosphorylation site"  
XX W0200042187-A1.  
XX 20-JUL-2000.  
PD

QY 65 RYLPEAYCLRGCLTGLFGEEDVFRSAPYMTVVLRRTPACAGRSVYTEAYVTIPVG 124  
|||||  
Db 92 RYLPEAYCLRGCLTGLFGEEDVFRSAPYMTVVLRRTPACAGRSVYTEAYVTIPVG 151  
|||||  
QY 125 CTCVPEPEKADSNSSIDKQAKLLGPNDA PGP 160  
|||||  
Db 152 CTCVPEPEKADSNSSIDKQAKLLGPNDA PGP 187  
|||||

Query Match 96.6%; Score 835; DB 21; Length 202;  
Best Local Similarity 98.7%; Pred. No. 2.1e-85;  
Matches 154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 5 ARVLSAFHHTLQIGPREQARNASCPAGRPADRRPPTNLRVSPWAYRISYDPARYP 64  
|||||  
Db 47 aagvlsafhtlqlgpreqarnascpagrpadrfrfppcnlrvspwayrissydparyp 106  
|||||  
QY 65 RYLPEAYCLRGCLTGLFGEEDVFRSAPYMTVVLRRTPACAGRSVYTEAYVTIPVG 124  
|||||

Db 107 RYLPEAYCLRGCLTGLFGEEDVFRSAPYMTVVLRRTPACAGRSVYTEAYVTIPVG 166  
QY 125 CTCVPEPEKADSNSSIDKQAKLLGPNDA PGP 160  
|||||  
Db 167 CTCVPEPEKADSNSSIDKQAKLLGPNDA PGP 202  
|||||

RESULT 8  
AAB07689  
ID AAB07689 standard; Protein; 202 AA.  
XX AC AAB07689;  
XX DT 07-NOV-2000 (first entry)  
XX DE A human interleukin-173 polypeptide.  
XX KW Interleukin; IL-171; cytokine; CTLA-8; IL-17; IL-175; IL-172; IL-173;  
XX IL-174; IL-176; IL-177; cell proliferation; cancer.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Peptide 1..17  
FT Protein /note= "signal peptide"  
FT Protein 18..202  
FT Modified-site /note= "mature protein"  
FT Modified-site 50  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 50..53  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 53  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 57..61  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 57..61  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 66..69  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 66..69  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 72..75  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 72..75  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 72  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 75  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 80  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 82  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 82..84  
FT Modified-site /note= "calcium phosphorylation site"  
FT Modified-site 113  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 113..116  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 113..116  
FT Modified-site /note= "CAMP protein kinase phosphorylation site"  
FT Modified-site 116  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 164..166  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 164..166  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 166..168  
FT Modified-site /note= "calcium phosphorylation site"  
XX W0200042187-A1.  
XX 20-JUL-2000.  
PD

XX 10-JAN-2000; 2000WO-US000005.  
XX 11-JAN-1999; 99US-0229402.  
XX (SCHE ) SCHERING CORP.  
XX Gorman DM, Bazan JF, Kastelein RA;  
XX WPI: 2000-476060/41.  
XX N-PSDB; AAA59155.  
XX New DNA sequence encoding a mammalian homolog of CTLA-8, designated  
PT interleukin-171 (IL-171), useful for recombinant production of IL-171  
PT which can be used for treating conditions associated with abnormal  
PT physiology or development -  
XX Disclosure; Page 16-17; 111pp; English.  
XX The present sequence represents an interleukin (IL)-173 polypeptide.  
CC It is a mammalian homologue of the cytokine designated CTLA-8 (also  
CC referred to as IL-17). The specification also describes homologues  
CC IL-171, IL-175, IL-172, IL-173, IL-174, IL-176, and IL-177. The DNA  
CC sequence encoding IL-171 is useful for identifying genes, mRNA and  
CC cDNA molecules which code for related or homologous proteins. The  
CC IL-171 protein, antibodies against IL-171, and compounds which have  
CC binding affinity to IL-171 are useful in treatment of conditions  
CC associated with abnormal physiology or development, including abnormal  
CC proliferation, e.g. cancerous conditions, or degenerative conditions.  
CC The IL-171 protein can be used in kits and assay methods for identifying  
CC compounds that selectively bind to IL-171.  
XX Sequence 202 AA;  
SQ  
Query Match 96.6%; Score 835; DB 21; Length 202;  
Best Local Similarity 98.7%; Pred. No. 2.1e-85;  
Matches 154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 5 ARAVLSAFHHTLQIGPREQARNASCPAGGRPADRRFPPTNLRSVSPWAYRISYDPAARYP 64  
Db 47 aagvlsafhtlqlgpreqarnascpaggrpadrrfpptnlrsvspwayrisydyaryp 106  
QY 65 RYLPEAYCLRCGCLTGLFGEDVFRFSAPVYMTVLLRTPACAGGRSVVTEAYVTIPVG 124  
Db 107 ryipeayclrcgcltglfgeedvfrfsapvymptvllrrtpacaggrsvvteayvtipvg 166  
QY 125 CTCVPEPEKDADSNSSIDKQAKLLGPNDAAPAGP 160  
Db 167 ctcvpepekadsinsidkqgklllgpnndapagp 202  
RESULT 9  
AAAY70653  
ID AAY70653 standard; Protein; 202 AA.  
AC AAY70653;  
XX 18-JUL-2000 (first entry)  
DT Human transforming growth factor beta-9, ztgbeta-9 protein.  
DE Human transforming growth factor beta-9; ztgbeta-9 protein.  
XX Human transforming growth factor beta-9; ztgbeta-9 protein.  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytostatic.  
XX Homo sapiens.  
OS Homo sapiens.  
FH Key  
FT Peptide  
FT /label= Signal\_peptide

Protein 16..202  
/note= "Mature transforming growth factor beta-9"  
WO200015798-A2.  
23-MAR-2000.  
17-SEP-1999; 99WO-US21677.  
17-SEP-1998; 98US-0154817.  
(ZYMO ) ZYMOGENETICS INC.  
Presnell SR, Taft DW, Foley KP;  
WPI: 2000-271436/23.  
N-PSDB; AAZ52195.  
Polynucleotides encoding a novel transforming growth factor beta-9  
antiproliferative agent -  
Claim 6; Page 83; 97pp; English.  
The present sequence is the human transforming growth factor beta-9,  
designated ztgf beta-9. Human ztgf beta-9 was isolated from an arrayed  
pituitary gland cDNA plasmid library by PCR screening. This can be used  
to treat a variety of neurodegenerative diseases such as amyotrophic  
lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
Parkinson's disease and peripheral neuropathies, or demyelinating  
diseases including multiple sclerosis. ztgf beta-9 peptides have  
antiviral activity and may also be used to regulate the proliferation,  
differentiation and apoptosis of neurons, glial cells, lymphocytes,  
hematopoietic cells and stromal cells.  
SQ Sequence 202 AA;  
Query Match 96.6%; Score 835; DB 21; Length 202;  
Best Local Similarity 98.7%; Pred. No. 2.1e-85;  
Matches 154; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 5 ARAVLSAFHHTLQIGPREQARNASCPAGGRPADRRFPPTNLRSVSPWAYRISYDPAARYP 64  
Db 47 aagvlsafhtlqlgpreqarnascpaggrpadrrfpptnlrsvspwayrisydyaryp 106  
QY 65 RYLPEAYCLRCGCLTGLFGEDVFRFSAPVYMTVLLRTPACAGGRSVVTEAYVTIPVG 124  
Db 107 ryipeayclrcgcltglfgeedvfrfsapvymptvllrrtpacaggrsvvteayvtipvg 166  
QY 125 CTCVPEPEKDADSNSSIDKQAKLLGPNDAAPAGP 160  
Db 167 ctcvpepekadsinsidkqgklllgpnndapagp 202  
RESULT 10  
AAAY70662  
ID AAY70662 standard; Protein; 209 AA.  
XX AAY70662;  
AC AAY70662;  
XX 18-JUL-2000 (first entry)  
DT Human transforming growth factor beta-9, ztgbeta-9 variant protein.  
DE Human transforming growth factor beta-9; ztgbeta-9 protein.  
XX Human transforming growth factor beta-9; ztgbeta-9 protein.  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytostatic.  
XX Homo sapiens.  
OS Homo sapiens.





Db 17 aagvisafhtlqlgpreqarnascpagrpdrfrftptnlrsvspwayisdydparyp 76  
QY 65 RYLPEAYCLRGCTGLTGLFGEEDVFRFSAPVYMPVTVLRRTPACAGGRSVYTEAYVTIPVG 124  
Db 77 rylpeayclrgcltglfgeedvfrfsapvymptvtvlrrtpacaggrsvyteayvtipvg 136  
QY 125 CTCVPEPEKDADSDIN 139  
Db 137 ctcvpepekdadsin 151  
RESULT 14  
AA070658  
ID AAY70658 standard; Protein; 183 AA.  
XX  
AC AAY70658;  
XX  
DT 18-JUL-2000 (first entry)  
XX  
DE Mature murine transforming growth factor beta-9, Ztgfbeta-9 protein.  
XX  
KW Murine transforming growth factor beta-9; Ztgf beta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytostatic.  
XX  
OS Mus sp.  
XX  
XX - WO200015798-A2.  
XX  
PD 23-MAR-2000.  
XX  
PF 17-SEP-1999; 99WO-0521677.  
XX  
PR 17-SEP-1998; 98US-0154817.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
XX Presnell SR, Taft DW, Foley KP;  
PI  
DR WPI; 2000-271436/23.  
DR N-PSDB; AA052198.  
XX  
PT Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated Ztgf beta-9, useful as an antiviral and  
PT antiproliferative agent  
XX  
PS Claim 6; Page 89; 97pp; English.  
XX  
CC The present sequence is the mature murine transforming growth factor  
CC beta-9, designated Ztgf beta-9. The signal sequence extends from  
CC amino acid residues 1 to 22. Murine Ztgf beta-9 was found to be highly  
CC expressed in the HCL hypothalamic cell line. This can be used  
CC to treat a variety of neurodegenerative diseases such as amyotrophic  
CC lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
CC Parkinson's disease and peripheral neuropathies, or demyelinating  
CC diseases including multiple sclerosis. Ztgf beta-9 peptides have  
CC antiviral activity and may also be used to regulate the proliferation,  
CC differentiation and apoptosis of neurons, glial cells, lymphocytes,  
CC hematopoietic cells and stromal cells.  
XX  
SQ Sequence 183 AA;  
Query Match 81.8%; Score 707; DB 21; Length 183;  
Best Local Similarity 84.5%; Pred. No. 3.8e-71;  
Matches 131; Conservative 6; Mismatches 14; Indels 4; Gaps 1;  
QY 5 ARVLSAFHTLQIGPREQARNASCPAGRPDRFRFTPTNLRSVSPWAYISVDPARYP 64  
Db 32 aagvisafhtlqlgpreqarnascpagrpdrfrftptnlrsvspwayisdydparif 91

QY 65 RYLPEAYCLRGCTGLTGLFGEEDVFRFSAPVYMPVTVLRRTPACAGGRSVYTEAYVTIPVG 124  
Db 92 rylpeayclrgcltglfgeedvfrfstpvpavlrirtaacaggrsvyaehtytipvg 151  
QY 125 CTCVPEPEKDADSDINSSIDKQAKLLGPNDPAG 159  
Db 152 ctcvpepekdadsanssmd----killigpdrpag 182  
RESULT 15  
AA07597  
ID AAB07597 standard; Protein; 205 AA.  
XX  
AC AAB07597;  
XX  
DT 07-NOV-2000 (first entry)  
XX  
DE A murine interleukin (IL) 173 polypeptide.  
XX  
KW Interleukin; IL-17; CTLA-8; IL-170; IL-172; IL-173; IL-174; IL-176;  
KW IL-177; IL-171; cell proliferation; cancer.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..24  
FT Protein /note= "signal peptide"  
FT Protein 25..205  
FT Modified-site /note= "mature protein"  
FT Modified-site 50  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 51..53  
FT Modified-site /note= "N-glycosylation site"  
FT Modified-site 53  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 57..61  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 80  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 82  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 82..84  
FT Modified-site /note= "calcium phosphorylation site"  
FT Modified-site 101..105  
FT Modified-site /note= "myristoylation site"  
FT Modified-site 113  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 116  
FT Modified-site /note= "phosphorylation site"  
FT Modified-site 164..166  
FT Modified-site /note= "N-glycosylation site"  
FT Modified-site 166..168  
FT Modified-site /note= "calcium phosphorylation site"  
XX  
PN WO200042188-A2.  
XX  
XX 20-JUL-2000.  
XX  
XX 10-JAN-2000; 2000WO-US000006.  
XX  
XX 11-JAN-1999; 99US-0228822.  
XX  
XX (SCHE ) SCHERING CORP.  
XX  
XX Gorman DM, Bazan JF, Kastelein RA;  
XX WPI; 2000-466130/40.  
XX N-PSDB; AAA58985.  
XX  
XX New isolated polynucleotide encoding a mammalian Interleukin-17 like  
XX protein used to identify genes for homologous proteins  
XX



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2002, 12:37:23 ; Search time 14.53 Seconds  
(without alignments)  
838.811 Million cell updates/sec

Title: US-09-731-816-4

Perfect score: 864

Sequence: 1 NSARARAVLSAFHHTLQLGP.....SIDKQAKLLGLGNDAPAGP 160

Scoring table: BLOSUM62  
Gapop 10.0 ; Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_68.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113.5	13.1	147	JC4628	cytotoxic T-lympho
2	108.5	12.6	151	B45351	immediate-early pr
3	108	12.5	150	T49623	cytotoxic T-lympho
4	99	11.5	148	T21334	hypothetical prote
5	83.5	9.7	226	T27843	hypothetical prote
6	77	8.9	354	T27099	hypothetical prote
7	77	8.9	603	S28941	coagulation factor
8	76	8.8	242	B81719	conserved hypothet
9	75.5	8.7	509	A84663	probable beta-keto
10	75.5	8.7	610	T44254	thiamin biosynthes
11	75	8.7	178	D72281	hypothetical prote
12	75	8.7	242	A71568	hypothetical prote
13	75	8.7	362	S22395	fetuin precursor -
14	74.5	8.6	181	T68674	cyclin-dependent k
15	74.5	8.6	1273	T38252	hypothetical prote
16	74.5	8.6	1896	1 RNFF2L	DNA-directed RNA p
17	74	8.6	495	T09993	diaminopimelate de
18	73	8.4	2397	1 A55535	transcan precursor
19	71.5	8.3	789	2 A39564	transcription repr
20	71	8.2	464	T22653	hypothetical prote
21	71	8.2	998	2 G83022	probable two-compo
22	70.5	8.2	181	T54380	cyclin-dependent k
23	70.5	8.2	256	T47860	transcription fact
24	70.5	8.2	265	D70778	hypothetical prote
25	70.5	8.2	479	1 A42241	glycine hydroxymet
26	70.5	8.2	872	T47564	probable ATP-depen
27	70.5	8.2	2115	2 S38480	nonstructural prot
28	70	8.1	206	2 T36643	probable integral
29	70	8.1	816	2 C69493	hypothetical prote

30	69.5	8.0	178	2 F96530	hypothetical prote
31	69	8.0	162	2 T32515	hypothetical prote
32	69	8.0	292	2 T03122	hypothetical prote
33	69	8.0	802	2 T24293	hypothetical prote
34	69	8.0	949	2 T24294	hypothetical prote
35	68.5	7.9	262	2 F75093	hypothetical prote
36	68.5	7.9	342	2 A56552	homeotic protein H
37	68	7.9	173	2 E71017	probable NADH-ubiq
38	68	7.9	298	2 A41230	DNA-3-methyladenin
39	68	7.9	528	2 T00951	probable 3-oxoacyl
40	68	7.9	585	2 A83020	probable carbamoyl
41	68	7.9	990	2 T14756	hypothetical prote
42	68	7.9	2205	1 MNWVRN	nonstructural poly
43	67.5	7.8	298	2 A32872	myogenic factor CM
44	67.5	7.8	426	2 A42360	cellulase (EC 3.2.
45	67.5	7.8	453	2 T01114	hypothetical prote

ALIGNMENTS

RESULT 1

JC4628

cytotoxic T-lymphocyte-associated antigen 8 precursor - mouse

N:Alternate names: CTLA8 protein

C:Species: Mus musculus (house mouse)

C:Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 26-Aug-1999

C:Accession: JC4628

R:Yao, Z.; Timour, M.; Painter, S.; Fanslow, W.; Spriggs, M.

Gene 168, 223-225, 1996

A:Title: Complete nucleotide sequence of the mouse CTLA8 gene.

A:Reference number: JC4628; MUID:96194901

A:Accession: JC4628

A:Molecule type: DNA

A:Residues: 1-147 <YAO>

A:Cross-references: GB:U35108; NID:g1244499; PIDN:AAA93253.1; PID:g1244500

C:Genetics:

A:Gene: ctla8

A:Introns: 69/2

C:Superfamily: saimiri herpesvirus immediate-early protein 2

C:Keywords: cytokine; glycoprotein; lymphocyte

F:1-14/Domain: signal sequence #status predicted <SIG>

F:15-147/Product: cytotoxic T-lymphocyte-associated antigen 8 #status predicted <NAT

F:60/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.1%; Score 113.5; DB 2; Length 147;  
Best Local Similarity 31.4%; Pred. No. 0.00033;  
Matches 37; Conservative 16; Mismatches 46; Indels 19; Gaps 6;

QY	32	GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLRGCLTGLFGEDVRFR	90
Db	47	GAKVSSR--RSDYLNRTSTPWLTHRNEDPDYPSVINEAQCRCRHCVCNA-EGKLDHNMN	103
QY	91	SAPVYMTPTVLLRRT-ACAGGRSVYETAYVTIPVGCTCPVEPEKDKASINSIDKGA	147
Db	104	SVLIQIQLVKREPESCP---FTFRVEKMLVGVCCTCV-----ASIVRQAA	147

RESULT 2

B45351

immediate-early protein 2 - saimirine herpesvirus 1 (strain 11)

N:Alternate names: hypothetical protein ORF13

C:Species: saimirine herpesvirus 1

A:Note: host Saimiri sciureus (common squirrel monkey)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999

C:Accession: B45351; D36807

R:Nicholas, J.; Smith, E.P.; Coles, L.; Honess, R.

Virology 179, 189-200, 1990

A:Title: Gene expression in cells infected with gammaherpesvirus saimiri: properties

A:Reference number: A45351; MUID:91021021

A:Accession: B45351

A:Molecule type: mRNA



Qy	143	DKQ5	143	
			:	:
Db	218	EGRG	221	
RESULT 8				
B81719	conserved hypothesis			
C:Species:	Chlamydia			
C:Date:	31-Mar-2000			
C:Accession:	B81719			
R:Read:	T.D.; Brun			
, C:	Dodson, R.;			
Nucleic Acids Res.				
A:Title:	Genome se			
A:Reference number:	8			
A:Accession:	B81719			
A>Status:	prelimin			
A:Molecule type:	D			
A:Residues:	1-242			
A:Cross-references				
A:Experimental sou				
C:Genetics:				
A:Gene:	TC0284			
Query Match				
Best Local Simil				
Matches 21; C				
Qy	68	PEAYCLCR		
		:	:	:
Db	161	PEPHCNCL		
Qy	116	-----		
Db	209	PLNPEQF		
RESULT 9				
A84663	probable beta-keto			
C:Species:	Arabido			
C:Date:	02-Feb-2000			
C:Accession:	A84666			
R:Lin, X.;	Kaul, S			
m., D.;	Nierman			
Nature	402, 761-766			
A:Title:	Sequence			
A:Reference number				
A:Accession:	A84666			
A>Status:	prelimin			
A:Molecule type:	D			
A:Residues:	1-509			
A:Cross-references				
C:Genetics:				
A:Gene:	At2g26640			
A:Map position:	2			
Query Match				
Best Local Simil				
Matches 26; C				
Qy	75	RCGLTGLE		
		:	:	:
Db	124	RSKLTCGF		
Qy	134	DADSN-S		
Db	173	EAEVTFG		
		:	:	:

**RESULT 10**

T44254  
thiamin biosynthesis protein thic [imported] - Rhizobium etli plasmid b  
C:Species: Rhizobium etli  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
C:Accession: T44254  
R:Miranda-Rios, J.; Morera, C.; Taboada, H.; Davalos, A.; Encarnacion, S.; Mora, J.; Soler, J. Bacteriol. 179, 6887-6893, 1997  
A:Title: Expression of thiamin biosynthetic genes (thiCOGE) and production of symbiotic A:Reference number: 222737; MUID:98037482  
A:Accession: T44254  
A>Status: preliminary; translated from GB/EMBL/DDBBJ  
A:Molecule type: DNA  
A:Residues: 1-610 <MIR>  
A:Cross-references: EMBL:AF004408; NID:g2627325; PIDN:AAC45972.1; PID:g2627326  
A:Experimental source: strain CE3  
C:Genetics:  
A:Gene: thiC  
A:Genome: plasmid b  
C:Function:  
A:Description: Involved in the biosynthesis of the pyrimidine moiety of thiamin  
C:Superfamily: thiamin biosynthesis protein thic

Query Match 8.7%; Score 75.5; DB 2; Length 610;  
Best Local Similarity 29.9%; Pred. No. 9.9;  
Matches 38; Conservative 14; Mismatches 54; Indels 21; Gaps 7;

Qy 30 PAGGRPADRRFRPPTNLRSVSPW---AYRISYDPARYPRYLPEAYCLRCGLTGLFGE-- 84  
Db | | | | | | | | | | | | | | | | | | | | | :  
| | | | | | | | | | | | | | | | | | | | | :  
44 PTSCEP-----PVTVDSSGPTYDPAHVLSID-AGLPR-LRESWIKARGDVESYDRIV 95  
| | | | | | | | | | | | | | | | | | | | | :  
Qy 85 --EDVFRSAPVMPTVVLTTPACA-GGRSVYTEAVVTIPVGCTCPVEPEKDADSINS 141  
| | | | | | | | | | | | | | | | | | | | | :  
Db 96 KPEDNGEATGERLTPEEPVRNTPLKAKAGRAVTLQAYAR-----AGITVPMEEFIAIREN 150  
| | | | | | | | | | | | | | | | | | | | | :  
Qy 142 IDKOGAK 148  
| | | | | | | | | | | | | | | | | | | | | :  
Db 151 LGROAAK 157

**RESULT 11**

D72281  
hypothetical protein TM1214 - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: D72281  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence  
A:Reference number: A72200; MUID:99287316  
A:Accession: D72281  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-178 <ARN>  
A:Cross-references: GB:AE001778; GB:AE000512; NID:g4981757; PIDN:AAD36289.1; PID:g498178  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TW1214  
C:Superfamily: psbG protein

Query Match 8.7%; Score 75; DB 2; Length 178;  
Best Local Similarity 23.1%; Pred. No. 2.9;  
Matches 34; Conservative 12; Mismatches 35; Indels 66; Gaps 7;

Qy 45 NLRSVSPWAVRI SYDPA RPYRLPEAYCLRCGC-----LTGLFGEE DVRFSA PV-- 94  
||||| | | | | | | | | | | | | | | | | | | | | | :  
Db 13 NLRKSRTIWLHL-----YCTGGCAVELPPSMTSRFDWE--RFGIAPMAT 53

Qy 95 -----YMTVTVLRRF-----PACAGGRSVVTEAYVT----- 120  
 Db 54 PROADILLITGLNTKRLRRVIYTEQMPDKYVVGFGSCTINGGIYFDSYATNRLDYY 113  
 Qy 121 IPVG---CTCVPEPEKDDSNSSIDK 144  
 Db 114 IPVDVYTAGCPRPEAILEAFNYLMEK 140  
 RESULT 12  
 A71568  
 hypothetical protein CT016 - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
 C:Accession: A71568  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchel  
 Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t  
 A:Reference number: A71570; MUID:9900809  
 A:Accession: A71568  
 A:Status: preliminary  
 A:Status: type: DNA  
 A:Molecule type: DNA  
 A:Residues: 1-242 <ARN>  
 A:Cross-references: GB:AE001276; GB:AE001273; NID:g3328399; PIDN:AAC67606.1; PID:g3332  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: CT016  
 Query Match 8.7%; Score 75; DB 2; Length 242;  
 Best Local Similarity 25.0%; Pred. No. 4.1;  
 Matches 20; Conservative 7; Mismatches 21; Indels 32; Gaps 3;  
 Qy 68 PEAVCLRCGLTGL-----FGEDVFRSPVYMTFVLLRTPACAGGRSVYT--- 115  
 Db 161 PEPHCNLCIOIGRATVEEDAGVSDLDLFRSWDI-----SOSGEKKMYTVD 208  
 Qy 116 -----EAYVTIPVGCTC 127  
 Db 209 PLNPEQFNVLGTPIGCTC 228  
 RESULT 13  
 S22395  
 fetuin precursor - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 21-Jul-2000  
 C:Accession: S22395; S22142  
 R:Brown, W.M.; Driedglewska, K.M.; Saunders, N.R.; Christie, D.L.; Nawratil, P.; Mue  
 Eur J. Biochem. 205, 321-331, 1992  
 A:Title: The nucleotide and deduced amino acid structures of sheep and pig fetuin. Co  
 A:Reference number: S22394; MUID:92209519  
 A:Accession: S22395  
 A:Molecule type: mRNA  
 A:Residues: 1-362 <BRW>  
 A:Cross-references: EMBL:X56021; NID:g2104; PIDN:CAA39498.1; PID:g3980229  
 C:Superfamily: alpha-2-HS-glycoprotein; cystatin homology  
 C:Keywords: calcium binding; EF hand; glycoprotein  
 F:1-15/Domain: signal sequence (fragment) #status predicted <SIG>  
 F:16-362/Product: fetuin #status predicted <MAT>  
 F:20-134/Domain: cystatin homology <CY1>  
 F:143-249/Domain: cystatin homology <CY2>  
 F:96, 153, 173/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 8.7%; Score 75; DB 2; Length 362;  
 Best Local Similarity 27.6%; Pred. No. 6.3;  
 Matches 27; Conservative 11; Mismatches 42; Indels 18; Gaps 3;

```

Query Match      8.7%; Score 75; DB 2; Length 362;
Best Local Similarity 27.6%; Pred. No. 6.3;
Matches 27; Conservative 11; Mismatches 42; Indels 18; Gaps

Qy 57 SYDPARYPRVLPEAYCLRCGLTGLFGEEDVFRSPAPVPMPTVVLRRTPACAGRSVYTE 116
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 210 AYSTPKNLIAVEKGYGCKGTVAKNVEDAVTCTVFOTVPOVLOPAGA----- 261

```

QY 117 AYVTIPVGCVCPEPEKDAKSINSI-DKOGAKLLGP 153  
 Db 262 -----DAGAT-----PVVDAATASPLADVPAASLVGVP 290  
 RESULT 14  
 I68674  
 cyclin-dependent kinase - human (fragment)  
 N:Alternate names: probable DNA synthesis inhibitor  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 01-Dec-2000  
 C:Accession: I68674; A9437; I53412; S39357  
 R:Mousses, S.; Ozcelik, H.; Lee, P.D.; Malkin, D.; Bull, S.B.; Andrulis, I.L.  
 Hum. Mol. Genet. 4, 1089-1092, 1995  
 A:Title: Two variants of the CIP1/WAF1 gene occur together and are associated with human  
 A:Reference number: I54380; MUID:95384154  
 A:Accession: I68674  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-181 <RES>  
 A:Cross-references: GB:L25610; NID:9425142; PIDN:AAAB59560.1; PID:g986879  
 R:Harper, J.W.; Adams, G.R.; Wei, N.; Keyomarsi, K.; Elledge, S.J.  
 Cell 75, 805-816, 1993  
 A:Title: The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent  
 A:Reference number: A49437; MUID:94061996  
 A:Accession: A49437  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 18-181 <RE3>  
 A:Cross-references: GB:L25610; NID:9425142; PIDN:AAAL6109.1; PID:g425143  
 R:Noda, A.; Ning, Y.; Venable, S.F.; Pereira-Smith, O.M.; Smith, J.R.  
 Exp. Cell Res. 211, 90-98, 1994  
 A:Title: Cloning of senescent cell-derived inhibitors of DNA synthesis using an express  
 A:Reference number: I53412; MUID:94170884  
 A:Accession: I53412  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 18-181 <RE2>  
 A:Cross-references: GB:L26165; NID:g418017; PIDN:AAAL19811.1; PID:g433742  
 R:Xiong, Y.; Hannon, G.J.; Zhang, H.; Casso, D.; Kobayashi, R.; Beach, D.  
 Nature 366, 701-704, 1993  
 A:Title: p21 is a universal inhibitor of cyclin kinases.  
 A:Reference number: S39357; MUID:94081955  
 A:Accession: S39357  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 18-181 <XIO>  
 A:Cross-references: GB:S67388; NID:g453134; PIDN:AAAB29246.1; PID:g453135  
 C:Genetics:  
 A:Gene: CIP1/WAF1

Query Match 8.6%; Score 74.5; DB 2; Length 181;  
 Best Local Similarity 24.6%; Pred. No. 3.4;  
 Matches 43; Conservative 10; Mismatches 73; Indels 49; Gaps 6;  
 QY 3 ARARAVLSAFHTLQLGPREQARNASCPAGGRPADRRFRPTNLRVSPPHAYRISYDPA 62  
 Db 4 ARIRRGTEALRGGMSEPAQGVQNPQNC---GSKACRLFGPVDSEQLS----- 48  
 QY 63 YPRYLPEAYCLRCGC-----LTFLEGEEDVFRFSAP-VYMPVTVLLRRT 104  
 Db 49 -----RCDALMACCIQEARERWNFDVETPTLEGDFANERVGLGLPKLYLPTGPRGR 103  
 QY 105 PACAGGRSVVT-----EAYVTIPVGCVCPEPEKDAK-SINSSIDKQAK 148  
 Db 104 DELGGRRPGTSPALLQGTAEEDVDLSLCTLYPSRGEQAEQSGPGGDSQGRK 158

RESULT 15  
 T38292  
 hypothetical protein SPAC23E2.02 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T38292  
 R:Stelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, January 1996  
 A:Reference number: Z21784  
 A:Accession: T38292  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1273 <SKE>  
 A:Cross-references: EMBL:Z68887; PIDN:CAA93114.1; GSPDB:GN00066; SPDB:SPAC23E2.02  
 A:Experimental source: strain 972h-; cosmid c23E2  
 C:Genetics:  
 A:Gene: SPDB:SPAC23E2.02  
 A:Map position: 1  
 A:Introns: 8/2; 862/3

Query Match 8.6%; Score 74.5; DB 2; Length 1273;  
 Best Local Similarity 25.6%; Pred. No. 27;  
 Matches 22; Conservative 15; Mismatches 48; Indels 1; Gaps 1;  
 QY 73 LCRGCLTGLFGEEDVFRFSAPVYMPVTVLLRTPACAGGRSVYTEAYVTIPVGCVCPEPE 132  
 Db 520 ICARQLTGLFSQYSSSFLSKNELPPKVIILEAKERTGGR-IYSRALPVVSHTSATQINHHT 578  
 QY 133 KDADSINSSIDKQAKLLLPNDAPA 158  
 Db 579 SNSNSISNSTSLAPKDVTDPSHIPS 604

Search completed: April 14, 2002, 12:38:08  
 Job time: 45 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2002, 12:37:48 ; Search time 11.74 Seconds  
(without alignments)  
499.691 Million cell updates/sec

Title: us-09-731-816-4

Perfect score: 864

Sequence: 1 NSARAVLSAFHHTLQLGP.....SIDKQAKLLGPNDAPAP 160

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135.5	15.7	155	1 IL17_HUMAN	Q16552 homo sapien
2	113.5	13.1	158	1 IL17_MOUSE	Q62386 mus musculus
3	108.5	12.6	151	1 VG13_HSVSA	P24916 herpesvirus
4	108	12.5	150	1 IL17_RAT	Q61453 rattus norv
5	78.5	9.1	1075	1 NFK3_MOUSE	P97305 mus musculus
6	77	8.9	603	1 FAL2_CAVPO	Q04962 cavia porce
7	75.5	8.7	425	1 Y450_HUMAN	O75038 homo sapien
8	75.5	8.7	610	1 THIC_RH1ET	O34291 rhizobium e
9	75	8.7	362	1 A2HS_PIG	P29700 sus scrofa
10	74.5	8.6	1273	1 YAR2_SCHPO	Q10135 schizosacch
11	74.5	8.6	1887	1 RPBL_DROME	P04052 drosophila
12	74	8.6	472	1 DCDA_MYCLE	Q50140 mycobacteri
13	73.5	8.5	1033	1 CR2_HUMAN	P20023 homo sapien
14	73	8.4	3358	1 PGCV_MOUSE	Q62059 mus musculus
15	71.5	8.3	164	1 CDN1_HUMAN	P38936 homo sapien
16	70.5	8.2	265	1 YM40_MYCTU	Q10522 mycobacteri
17	70.5	8.2	479	1 GLYC_NEUCR	P34898 neurospora
18	69.5	8.0	324	1 TTP_BOVIN	P53781 bos taurus
19	69	8.0	397	1 A2BP_HUMAN	Q9nwb1 homo sapien
20	68.5	7.9	342	1 HXCA_MOUSE	P31257 mus musculus
21	68	7.9	2205	1 POLN_RUBVT	P13889 rubella vir
22	67.5	7.8	298	1 MYOD_CHICK	P16075 gallus gall
23	67.5	7.8	384	1 DUS9_HUMAN	Q99956 homo sapien
24	67.5	7.8	591	1 UL49_EBV	P14347 Epstein-bar
25	67	7.8	277	1 HXDB_NOTVI	P31263 notophthalm
26	67	7.8	2273	1 HFAL_YEAST	P32874 saccharomyc
27	67	7.8	4393	1 PGBM_HUMAN	P98160 homo sapien
28	66.5	7.7	1117	1 TCF8_MOUSE	Q64318 mus musculus
29	66.5	7.7	2339	1 CCAB_RABIT	Q05152 oryctolagus
30	66	7.6	489	1 VHS_HSV1	P10225 herpes simp
31	66	7.6	676	1 DNLI_THETH	P26996 thermus aqu
32	65.5	7.6	170	1 PLGF_HUMAN	P49763 homo sapien
33	65.5	7.6	369	1 CHEB_RHOSH	O33558 rhodobacter

## ALIGNMENTS

RESULT 1	IL17_HUMAN	STANDARD;	PRT;	155 AA.
34	65.5	7.6	404	1 VE2_HPV60
35	65.5	7.6	500	1 STCL_EMENI
36	65	7.5	337	1 YJN_ECOLI
37	65	7.5	376	1 METK_AQUAE
38	65	7.5	459	1 Y976_TREPA
39	65	7.5	802	1 ACSE_ACEXY
40	65	7.5	1106	1 GLI1_HUMAN
41	64.5	7.5	285	1 HXA4_MOUSE
42	64.5	7.5	342	1 HXCA_HUMAN
43	64.5	7.5	374	1 RLIG_BPT4
44	64	7.4	286	1 PYRL_SYNEL
45	64	7.4	396	1 A2BP_MOUSE
Q80944	human papil			
Q00707	emericeilla			
P39400	escherichia			
O67222	aquifex ae			
O83941	treponema p			
P37716	acetobacter			
P08151	homo sapien			
P06798	mus musculu			
Q9ayd6	homo sapien			
P00971	bacterioph			
P50034	synechococ			
Q9jj43	mus musculu			
AC	Q16552;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED			
DE	ANTIGEN 8) (CTLA-8).			
GN	IL17 OR CTLA8.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96281911; PubMed=8676080;			
RA	Fossiez F., Djossou O., Chomarat P., Flores-Romo L., Ait-Yahia S.,			
RA	Maat C., Pin J.-J., Garrone P., Garcia E., Saeland S., Blanchard D.,			
RA	Gaillard C., Das Mahapatra B., Rouvier E., Golstein P., Banchereau J.,			
RA	Lebecqec S.			
RT	"T cell interleukin-17 induces stromal cells to produce			
RT	proinflammatory and hematopoietic cytokines."			
RL	J. Exp. Med. 183:2593-2603(1996).			
RL	[2]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=T-cell;			
RX	MEDLINE=96094436; PubMed=7499828;			
RA	Yao Z., Painter S.L., Fanslow W.C., Ulrich D., Macduff B.M.,			
RA	Spriggs M.K., Armitage R.J.;			
RT	"Human IL-17: a novel cytokine derived from T cells."			
RL	J. Immunol. 155:5483-5486(1995).			
CC	-1- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND			
CC	HEMATOPOIETIC CYTOKINES. ENHANCES THE SURFACE EXPRESSION OF THE			
CC	INTRACELLULAR ADHESION MOLECULE-1 (ICAM-1) IN FIBROBLASTS.			
CC	-1- SUBUNIT: HOMODIMER.			
CC	-1- SUBCELLULAR LOCATION: SECRETED.			
CC	-1- TISSUE SPECIFICITY: RESTRICTED TO ACTIVATED MEMORY T-CELLS.			
CC	-1- PTM: FOUND BOTH IN GLYCOSYLATED AND NONGLYCOSYLATED FORMS.			
CC	-1- SIMILARITY: STRONG. TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO			
CC	HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; Z58820; CAA91233.1; -			
DR	EMBL; U32659; AAC50341.1; -			
DR	MM; 603149; -			
DR	1 23			
KW	Cytokine; Glycoprotein; T-cell; Antigen; Signal.			
FT	CHAIN 1 23			
FT	24 155			
FT	INTERLEUKIN-17.			



```
FT CARBOHYD 64 64 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 151 AA; 17180 MW; 53BEDDE4206C6432 CRC64;

Query Match 12.68; Score 108.5; DB 1; Length 151;
Best Local Similarity 31.48; Pred. No. 0.00021;
Matches 27; Conservative 11; Mismatches 37; Indels 11; Gaps 3;

QY 47 RVSFPMAYRISYDPRYPRLPEAYCICRGCLTGLFGEDVRFRSAPVYMTVVLRRTPA 106
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 65 RSTSPWTLHRNEDQDRPVSIVWEAKRVLGCVNA-DGNVDYHMNSVPIQOEILVVRK--- 120

QY 107 CAGGRSVYTEAY----VTIPVGCTCV 128
   | : : : |||||
DB 121 ---GHQPCPNFSRLEKMLVTGCTCV 143

RESULT 4
IL17_RAT STANDARD; PRT; 150 AA.
AC Q61453;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED
DE ANTIGEN 8) (CTLA-8).
GN IL17 OR CTLA8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93294300; PubMed=8390535;
RA Rouvier E., Luciani M.F., Mattei M.-G., Denizot F., Golstein P.;
RT "CTLA-8, cloned from an activated T cell, bearing AU-rich messenger
RT RNA instability sequences, and homologous to a herpesvirus saimiri
RT gene.";
RL J. Immunol. 150:5445-5456(1993).
RN [2]
RP ORGANISM IDENTIFICATION.
RX MEDLINE=96194901; PubMed=8654948;
RA Yao Z., Timour M., Painter S., Fanslow W., Spriggs M.K.;
RT "Complete nucleotide sequence of the mouse CTLA8 gene.";
RL Gene 168:223-225(1996).
RN [3]
RP SEQUENCE FROM N.A., AND ORGANISM IDENTIFICATION.
RX MEDLINE=97031826; PubMed=8877732;
RA Kennedy J., Rossi D.L., Zurawski S.M., Vega F. Jr., Kastelein R.A.,
RA Wagner J.L., Hannum C.H., Zlotnik A.;
RT "Mouse IL-17: a cytokine preferentially expressed by alpha beta TCR +
RT CD4-CD8-T cells.";
RL J. Interferon Cytokine Res. 16:611-617(1996).
CC -1- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND
CC HEMATOPOIETIC CYTOKINES (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: STRONG. TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO
CC HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE FROM MOUSE BUT, ON
CC THE BASIS OF SUBSEQUENT WORK (REF.2 AND REF.3), HAS BEEN SHOWN TO
CC BE OF RAT ORIGIN.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L13839; AAA37490.1;
KW Cytokine; Glycoprotein; T-cell; Antigen; Signal.
```

```
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 150 INTERLEUKIN-17.
FT CARBOHYD 63 63 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 46 46 I -> L (IN REF.3).
SQ SEQUENCE 150 AA; 16876 MW; EF13F33EDF9D689F CRC64;

Query Match 12.58; Score 108; DB 1; Length 150;
Best Local Similarity 34.48; Pred. No. 0.00023;
Matches 31; Conservative 11; Mismatches 42; Indels 6; Gaps 4;

QY 41 RPPNML-RVSPMAYRISYDPRYPRLPEAYCICRGCLTGLFGEDVRFRSAPVYMTV 99
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 57 RPSDYLNRSTSPWTLHRNEDQDRPVSIVWEAQRHRCVNA-EGKLDHMHNSVLIQOEIL 115

QY 100 VLRRTP-ACAGGRSVYTEAYVTIPVGCTCV 128
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 116 VLKREPKCP---FTFRVERKMLVGVGCTCV 142

RESULT 5
NFC3_MOUSE STANDARD; PRT; 1075 AA.
AC P97305; Q60896;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NUCLEAR FACTOR OF ACTIVATED T-CELLS, CYTOPLASMIC 3 (T CELL
DE TRANSCRIPTION FACTOR NFAT4) (NF-ATC3) (NF-AT4) (NFATX).
GN NFATC3 OR NFAT4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM X1).
RX TISSUE=Thymus;
RX MEDLINE=95378239; PubMed=7650004;
RA HO S.N., Thomas D.J., Timmerman L.A., Li X., Francke U.,
RA Crabtree G.R.;
RT "NFATc3, a lymphoid-specific NFATc family member that is
RT calcium-regulated and exhibits distinct DNA binding specificity.";
RL J. Biol. Chem. 270:19898-19907(1995).
RN [2]
RP SEQUENCE OF 12-1075 FROM N.A. (ISOFORMS X1; X2 AND DELTA-X).
RX TISSUE=Thymic lymphoma;
RX MEDLINE=97170074; PubMed=9017603;
RA Liu J., Koyano-Nakagawa N., Amasaki Y., Saito-Obara F., Ikeuchi T.,
RA Inai S.-I., Takano T., Arai N., Yokota T., Arai K.-I.;
RT "Calcineurin-dependent nuclear translocation of a murine transcription
RT factor NFATc: molecular cloning and functional characterization.";
RL Mol. Biol. Cell 8:157-170(1997).
RN [3]
RP REVIEW.
RX MEDLINE=99189746; PubMed=10089876;
RA Crabtree G.R.;
RT "Generic signals and specific outcomes: signaling through Ca2+,
RT calcineurin, and NF-AT.";
RL Cell 96:611-614(1999).
CC -1- FUNCTION: PLAYS A ROLE IN THE INDUCIBLE EXPRESSION OF CYTOKINE
CC GENES IN T CELLS, ESPECIALLY IN THE INDUCTION OF THE IL-2 (BY
CC SIMILARITY).
CC -1- SUBUNIT: MEMBER OF THE MULTICOMPONENT NFATC TRANSCRIPTION COMPLEX
CC THAT CONSISTS OF AT LEAST TWO COMPONENTS, A PRE-EXISTING
CC CYTOPLASMIC COMPONENT NFATC2 AND AN INDUCIBLE NUCLEAR COMPONENT
CC NFATC1. OTHER MEMBERS SUCH AS NFATC4, NFATC3 OR MEMBERS OF THE
CC ACTIVATING PROTEIN-1 FAMILY, MAF, GATA4 AND CBP/300 CAN ALSO BIND
CC THE COMPLEX. NFATC PROTEINS BIND TO DNA AS MONOMERS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FOR THE PHOSPHORYLATED FORM AND
CC NUCLEAR AFTER ACTIVATION THAT IS CONTROLLED BY CALCINEURIN-
CC MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC IS THOUGHT
CC TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED
CC AND TRANSIENT CALCIUM SIGNALS. THE SUBCELLULAR LOCALIZATION OF
```

```

CC CC NPATC PLAY A KEY ROLE IN THE GENE TRANSCRIPTION.
CC CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; X1 (SHOWN HERE), X2 AND DELTA-X;
CC CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC CC -1- TISSUE SPECIFICITY: EXPRESSED IN THYMUS. WEAKLY EXPRESSED IN
CC CC MUSCLE, SPLEEN AND KIDNEY. ALSO EXPRESSED IN LYMPH NODE.
CC CC -1- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND
CC CC COOPERATIVE INTERACTIONS WITH AP1 FACTORS (BY SIMILARITY).
CC CC -1- PTM: PHOSPHORYLATED BY NFATC-KINASE; DEPHOSPHORYLATED BY
CC CC CALCINEURIN (BY SIMILARITY).
CC CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; D85612; BAA12833.1; -.
CC CC EMBL; U28807; AAA93249.1; -.
CC CC MGI; MGI:103296; Nfatc3.
CC CC InterPro; IPR002909; IPT_TIG.
CC CC InterPro; IPR000451; REL.
CC CC Pfam; PF00554; RHD; 1.
CC CC SMART; SM00429; IPT; 1.
CC CC PROSITE; PS01204; REL_1; FALSE_NEG.
CC CC PROSITE; PS0254; REL_2; 1.
CC CC Transcription regulation; Activator; Nuclear protein; DNA-binding;
CC CC Alternative splicing; Repeat; Phosphorylation.
CC CC DOMAIN 24 29 POLY-PRO.
CC CC FT DOMAIN 109 114 CALCINEURIN-BINDING.
CC CC FT DOMAIN 207 308 3 X SP REPEATS.
CC CC FT REPEAT 207 223 1.
CC CC FT REPEAT 236 252 2.
CC CC FT REPEAT 292 308 3.
CC CC FT DOMAIN 444 451 NUCLEAR LOCALIZATION SIGNAL.
CC CC FT DOMAIN 686 688 NUCLEAR LOCALIZATION SIGNAL.
CC CC FT DOMAIN 1031 1040 NUCLEAR EXPORT SIGNAL.
CC CC FT VARSPLIC 468 497 MISSING (IN ISOFORM DELTA-X).
CC CC FT VARSPLIC 1035 1075 VNEIGHDSQISVSQATEVMYRDTPLPGPASPDMLTSHSAH
CC CC -> DOLISLDEKHPGSGTSEKHNHSEFSCVPFWRI (IN
CC CC ISOFORM X2).
CC CC FT CONFLICT 12 34 LDKLVFGDGAPAPPVSGKPA -> PRRVLSVSQAQLPS
CC CC FT RTRPGPSDL (IN REF. 2).
CC CC FT CONFLICT 57 61 TTPIC -> NSSIG (IN REF. 2).
CC CC FT CONFLICT 68 68 P -> IQ (IN REF. 2).
CC CC FT CONFLICT 82 83 SH -> GY (IN REF. 2).
CC CC FT CONFLICT 90 93 CEIP -> GOIS (IN REF. 2).
CC CC FT CONFLICT 113 113 F -> I (IN REF. 2).
CC CC FT CONFLICT 121 124 QDEL -> HOGT (IN REF. 2).
CC CC FT CONFLICT 132 132 Q -> H (IN REF. 2).
CC CC FT CONFLICT 140 140 F -> Y (IN REF. 2).
CC CC FT CONFLICT 641 641 F -> R (IN REF. 2).
CC CC FT CONFLICT 646 646 V -> A (IN REF. 2).
CC CC FT CONFLICT 707 734 REDTDLPSVPSLPVPHSAQAQRPSSETG -> KEKTOICILQ
CC CC FT FHLCCLILPSALLRDR (IN REF. 2).
CC CC FT CONFLICT 746 746 G -> S (IN REF. 2).
CC CC FT CONFLICT 1075 1075 H -> Q (IN REF. 2).
CC CC SEQUENCE 1075 AA; 115450 MW; 4ED38C9AA6F452BB CRC64;

Query Match 9.18; Score 78.5; DB 1; Length 1075;
Best Local Similarity 22.08; Pred. No. 2.7;
Matches 37; Conservative 22; Mismatches 76; Indels 33; Gaps 5;

Oy 1 NSARARAVLSAFHHTLQLGPREQARNASCPA-----GGRPADRRFRP 42
Db 684 NGKRKKSQSRTFTVPLMKQEQREDTDLPSVPSLPVPHSAQAQRPSSETPHSDRAMSA 743
Oy 43 PTNLR-SVSPWAYRISDPARYPRYLPEAYCLRCGCTGLGEGEDVRFSPAPVTMPTVVL 101

```

```

Db 744 PGGLLCQVQP-----AYTSMVASTHLFQLQCRDEGA-----GKEQHIATSSVMHQFOVT 793
Oy 102 RRTFACAGGSRVYEAVTTPGTCVPEPEKADADSNSSIDKQAGAL 149
Db 794 PTSPIGSSYSIQTSMTY-----NGPTCLPVPNPASSQEFDPVLFQDDAAL 837

RESULT 6
FAL2_CAVPO STANDARD; PRT; 603 AA.
AC Q04962; 1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE COAGULATION FACTOR XII PRECURSOR (EC 3.4.21.38) (HAGEMAN FACTOR)
DE (HAF) (FRAGMENT).
GN F12.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC NCBI_TaxID=10141;
OX NCBI_TaxID=10141;
RN [1]
RC SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RP TISSUE=Liver;
RX MEDLINE=93003367; PubMed=1390917;
RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
RA Kambara T., Okabe H.;
RT "Primary structure of guinea-pig Hageman factor: sequence around the
RT cleavage site differs from the human molecule.";
RL Biochim. Biophys. Acta 1159:113-121(1992).
CC -1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -1- CATALYTIC ACTIVITY: CLEAVES SELECTIVELY ARG-|-ILE BONDS AND
CC ACTIVATES COAGULATION FACTORS VII AND XI.
CC -1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68615; CAA48600.1; -.
CC HSP: P00763; IDPO.
CC MEROPS; S01.211; -.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR000562; FN_Type_II.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Trypsin.
CC InterPro; IPR000083; fibronectin_type_1.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC ProDom; PD000995; FN_Type_II; 1.
CC SMART; SM00181; EGF; 2.

```

DR SMART; SM00058; FN1; 1.  
DR SMART; SM00059; FN2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.  
FT NON\_TER 1 18  
FT SIGNAL <1 18  
FT CHAIN 19 358 ALPHA-FACTOR XIIA HEAVY CHAIN.  
FT CHAIN 359 603 ALPHA-FACTOR XIIA LIGHT CHAIN.  
FT DOMAIN 46 87 FIBRONECTIN TYPE-II.  
FT DOMAIN 93 130 EGF-LIKE 1.  
FT DOMAIN 132 172 FIBRONECTIN TYPE-I.  
FT DOMAIN 173 209 EGF-LIKE 2.  
FT DOMAIN 216 294 KRINGLE.  
FT DOMAIN 312 342 PRO-RICH.  
FT DOMAIN 359 603 SERINE PROTEASE.  
FT ACT\_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 97 109 BY SIMILARITY.  
FT DISULFID 103 118 BY SIMILARITY.  
FT DISULFID 120 129 BY SIMILARITY.  
FT DISULFID 134 162 BY SIMILARITY.  
FT DISULFID 160 169 BY SIMILARITY.  
FT DISULFID 177 188 BY SIMILARITY.  
FT DISULFID 182 197 BY SIMILARITY.  
FT DISULFID 199 208 BY SIMILARITY.  
FT DISULFID 216 294 BY SIMILARITY.  
FT DISULFID 237 276 BY SIMILARITY.  
FT DISULFID 265 289 BY SIMILARITY.  
FT DISULFID 345 472 BY SIMILARITY.  
FT DISULFID 383 399 BY SIMILARITY.  
FT DISULFID 391 461 BY SIMILARITY.  
FT DISULFID 422 425 BY SIMILARITY.  
FT DISULFID 488 557 BY SIMILARITY.  
FT DISULFID 520 536 BY SIMILARITY.  
FT DISULFID 547 578 BY SIMILARITY.  
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9ED59 CRC64;

Query Match 8.9%; Score 77; DB 1; Length 603;

Best Local Similarity 21.2%; Pred. No. 2;  
Matches 39; Conservative 18; Mismatches 59; Indels 68; Gaps 8;

QY 12 FHHTLQLGPREQARNASCPAGGRADRRFRPT-NLRVSPWAYRISYDPAIPRY----- 66  
DB 57 YHH-----CIHKGGRGPRWCATTNFDDQDQWVCL--EPKVKDHCKSH 100  
QY 67 -----LPEAYCLRCGLTG-----LFGEEDVRRFRSAPV----- 94  
DB 101 NPCQGGICVNTLSSPHCLCLDHLTKHCQREKCFEPOLHRFFHENEIWFRTGPAVAKC 160  
QY 95 -----YMPVTWVLRRTACAGGRSVYTEAY--VTIPVGCVCVPEPKADSISSI 142  
DB 161 HCKGPDHCKQMHSEQCQTNCLNGRCLEVEGHLCDCPMGYT---GPFCDLDTTASY 217  
QY 143 DKQG 146  
DB 218 EGRG 221

RESULT 7  
Y450\_HUMAN  
ID Y450\_HUMAN STANDARD; PRT; 425 AA.  
AC OY5038;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL PROTEIN KIAA0450.  
GN KIAA0450.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=98116662; PubMed=9455484;  
RA Seki N., Ohira M., Nagase T., Ishikawa K.-I., Miyajima N.,  
RA Nakajima D., Nomura N., Ohara O.;  
RT "Characterization of cDNA clones in size-fractionated cDNA libraries  
from human brain.";  
RL DNA Res. 4:345-349(1997).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AB007919; BAA32295.1;  
KW Hypothetical protein.  
SQ SEQUENCE 425 AA; 43842 MW; 2A6D733CA149E665 CRC64;  
  
Query Match 8.7%; Score 75.5; DB 1; Length 425;  
Best Local Similarity 28.3%; Pred. No. 1.9;  
Matches 39; Conservative 13; Mismatches 43; Indels 43; Gaps 10;  
  
QY 20 PREQARNASCPAGGR-----PADRRFRPTNLRVSPWAYR-----ISVD----- 59  
DB 270 PRARADSLGAPCCGLDHPAIPGRSREAPK-----GPCAWRQGGSGMSDSSSDSP 323  
QY 60 --PARYPRYLPEAYCLRCGLTG---LFGE--EDVRRFRSAPV-----MPTVLRTP 105  
DB 324 GIPERSPRW-PEGACRQPGALOGEMSAFQAOKLEEIRSKS-PMFSAGKPLLPVVVLPHP 381  
QY 106 ACAG-GRSVYTEAVVTIP 122  
DB 382 GMAGPGSPAAASAWTVSP 399  
  
RESULT 8  
THIC\_RHET  
ID THIC\_RHET STANDARD; PRT; 610 AA.  
AC O34291;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE THIAMINE BIOSYNTHESIS PROTEIN THIC.  
GN THIC.  
OS Rhizobium etli.  
OG Plasmid pb.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=29449;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CE3;  
RX MEDLINE=98037482; PubMed=9371431;  
RA Miranda-Rios J., Morera C., Taboada H., Davalos A., Encarnacion S.,



DR InterPro: IPR002937; Amino\_oxidase.  
 DR InterPro: IPR000910; HMG\_12\_box.  
 DR Pfam: PF01593; Amino\_oxidase; 1.  
 DR SMART: SM00398; HMG; 1.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 1028 1048 POTENTIAL.  
 SQ SEQUENCE 1273 AA; 142489 MW; 3FEFF6433DF3C0B6 CRC64;

Query Match 8.6%; Score 74.5; DB 1; Length 1273;  
 Best Local Similarity 25.6%; Pred. No. 8.6;  
 Matches 22; Conservative 15; Mismatches 48; Indels 1; Gaps 1;

QY 73 LCRGLTGLFGEEDVRFSAPVYMTVVRTPACAGRSYTYRATYIPVGCVCVPEPE 132  
 :| | | | | : | | | | | : | | | | | : | | | | | :  
 Db 520 ICARQLTGLFQYSSFLSKNELPKVILEAKERTGGR-IYSRALPVSHTSATQINHHT 578

QY 133 KQADINSIDKQAKLLLPNDAPA 158  
 :| | | | | : | | | | | : | | | | | : | | | | | :  
 Db 579 SNSNSISNSTSLNPKDVTDPFSHP 604

RESULT 11  
 RPBL\_DROME  
 AC P04052; Q9VYX6;  
 DT 01-NOV-1986 (Rel. 03, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT (EC 2.7.7.6).  
 GN RPII215 OR CG1554.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89218930; PubMed=2496296;  
 RA Jocker R.S., Weeks J.R., Zehring W.A., Greenleaf A.L.;  
 RT "Analysis of the gene encoding the largest subunit of RNA polymerase  
 RT II in *Drosophila*."  
 RM Mol. Gen. Genet. 215:266-275(1989).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananathides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP SEQUENCE OF 1-27 FROM N.A.  
 RX MEDLINE=8709662; PubMed=3025586;  
 RA Searles L.L., Greenleaf A.L., Kemp W.E., Voelker R.A.;  
 RT "Sites of P element insertion and structures of P element deletions in  
 RT the 5' region of *Drosophila melanogaster* RPII215."  
 RL Mol. Cell. Biol. 6:3312-3319(1986).  
 RN [4]  
 RP SEQUENCE OF 1-472 FROM N.A.  
 RX MEDLINE=85282618; PubMed=2992806;  
 RA Biggs J., Searles L.L., Greenleaf A.L.;  
 RT "Structure of the eukaryotic transcription apparatus: features of the  
 RT gene for the largest subunit of *Drosophila* RNA polymerase II."  
 RL Cell 42:611-621(1985).  
 RN [5]  
 RP SEQUENCE OF 1441-1887 FROM N.A.  
 RX MEDLINE=88094402; PubMed=312024;  
 RA Allison L.A., Wong J.K.-C., Fitzpatrick V.D., Moyle M., Ingles C.J.;  
 RT "The C-terminal domain of the largest subunit of RNA polymerase II of  
 RT *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and mammals: a  
 RT conserved structure with an essential function."  
 RL Mol. Cell. Biol. 8:321-329(1988).  
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.  
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +  
 CC RNA(N).  
 CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO  
 CC TWELVE DIFFERENT POLYPEPTIDES. THE 215 KDA POLYPEPTIDE IS THE  
 CC LARGEST COMPONENT OF RNA POLYMERASE II.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.  
 CC -1- THE PHOSPHORYLATION ACTIVATES POL2.  
 CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE  
 CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA  
 CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE  
 CC III FOR 5S AND TRNA GENES.  
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M27431; AAA28868.1; -  
 CC EMBL: AE003486; AAF48057.1; -  
 CC EMBL: M14203; AAA28864.1; -  
 CC EMBL: M11798; AAA28863.1; -  
 CC EMBL: M19537; AAA28827.1; -  
 CC PIR: S04457; RNPF2L.  
 CC FlyBase: FBgn003277; RPII215.  
 CC InterPro: IPR000684; RNA\_polII\_repeat.  
 CC InterPro: IPR000722; RNA\_pol\_A.  
 CC InterPro: IPR002879; RNA\_pol\_A2.  
 CC Pfam: PF00623; RNA\_pol\_A; 1.  
 CC Pfam: PF01854; RNA\_pol\_A2; 1.  
 CC PROSITE: PS00115; RNA\_POL\_II\_REPEAT; 11.  
 CC Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;

```

KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT ZN_FING 67 83 C2H2-TYPE (POTENTIAL).
FT DOMAIN 1579 1881 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
FT CONFLICT 319 324 RAMOKS -> GYAKV (IN REF. 4).
FT CONFLICT 450 450 F -> G (IN REF. 4).
FT CONFLICT 455 458 TLHK -> RCTT (IN REF. 4).
FT CONFLICT 463 472 GHRVKLPWS -> VTGESVASST (IN REF. 4).
FT CONFLICT 741 741 R -> H (IN REF. 1).
FT CONFLICT 1485 1524 SMLGGAAMFICGGSTPMTPTWANCNTPRVFSPGPHV
      -> I (IN REF. 5).
FT CONFLICT 1506 1508 MTP -> ELDSA (IN REF. 1).
FT CONFLICT 1887 1887 D -> DVKGGRG (IN REF. 1).
SQ SEQUENCE 1887 AA; 209167 MW; 4EC68C7708A167A3 CRC64;

Query Match      8.6%; Score 74.5; DB 1; Length 1887;
Best Local Similarity 24.8%; Pred. No. 13;
Matches 37; Conservative 15; Mismatches 50; Indels 47; Gaps 9;

Qy 20 PREQARNASCAGGRPADRRFRPT-NLRVSPWAYRISYDPARY----PRYLPEAYCLC 74
Db 1581 PTSPNYTASSPGGASP---NYPSSPNYSPTSP-----LYASPRYASTPTNFNPQS----- 1628
Qy 75 RGCLTGLFGEEDVFRSPAPVYPTVLRRTTACAG-GRSVYT--EAYVTIPVGCTCVPEP 131
Db 1629 ----TGVPSSSGYSPTSPVYPTVQFQSSPSFAGSGSNISPGNAV-----SP 1673
Qy 132 EKDADSNSSIDKQAKLLGPNDAAPG 160
Db 1674 SSSNYSPTS-----PSYSPTSP 1690

RESULT 12
DCDA_MYCLE STANDARD; PRT; 472 AA.
AC Q50140; Q9CC78;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE -DIAMINOPIMELATE DECARBOXYLASE (EC 4.1.1.20) (DAP DECARBOXYLASE).
GN LYSA OR ML1128.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Mycobacterium.
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP Smith D.R., Robison K.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajadream M.-A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC CC -1- CATALYTIC ACTIVITY: MESO-2,6-DIAMINOHEPTANEDIOATE = L-LYSINE +
CO(2).
CC CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC CC -1- PATHWAY: LAST STEP IN THE BIOSYNTHESIS OF LYSINE.
CC CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGinine
DECARBOXYLASES.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U15186; AAA63102.1; ALT_INIT.
DR EMBL; AL583920; CAC31509.1; -.
DR Leproma; ML1128; -.
DR InterPro; IPR000183; Orn_DAP_Arg_decarbxylyse.
DR Pfam; PF00278; Orn_DAP_Arg_dec; 1.
DR PRINTS; PRO1179; ODADCRBXLASE.
DR PRINTS; PRO1181; DAPDCRBXLASE.
DR PROSITE; PS00878; ODR_DC_2.1; 1.
DR PROSITE; PS00879; ODR_DC_2.2; FALSE_NEG.
KW Lysine biosynthesis; Lyase; Decarboxylase; Pyridoxal phosphate;
KW Complete proteome.
FT BINDING 97 97 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 472 AA; 50277 MW; DB234E82D8274B39 CRC64;

Query Match      8.6%; Score 74; DB 1; Length 472;
Best Local Similarity 25.0%; Pred. No. 3.1;
Matches 37; Conservative 16; Mismatches 47; Indels 48; Gaps 8;

Qy 14 HTLQLGPR--EQARNASCAGGRPADRRFRPTNLRVSPWAYRISYDPARYPRYLPEAY 71
Db 4 HT--AGPRHAEKTRHTATPQVQPSDDLRLASNV----WPRNITRDET-----GVA 49
Qy 72 CLCRGCLTGLFG-----EDVFRSPAPVYPTVLRRTTACAGGRSVYETAYVTI 121
Db 50 CIAGNKLTDLAGEYGTPLFVIDEDDFRCREI-----AAAFGGGVNHYAAKAF 100
Qy 122 PVGCTCVPEPEKDADSNSSIDKQAKL 149
Db 101 ---CT-----ETARWIDEGLSL 115

RESULT 13
CR2_HUMAN
ID CR2_HUMAN STANDARD; PRT; 1033 AA.
AC P20023;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE COMPLEMENT RECEPTOR TYPE 2 PRECURSOR (CR2) (COMPLEMENT C3D RECEPTOR)
DE (EPSTEIN-BARR VIRUS RECEPTOR) (EBV RECEPTOR) (CD21 ANTIGEN).
DE CR2 OR C3DR.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89123277; PubMed=2563370;
RA Fujisaku A., Harley J.B., Frank M.B., Gruner B.A., Frazier B.,
RA Holers V.M.;
RT "Genomic organization and polymorphisms of the human C3d/Epstein-Barr
RT virus receptor.";
RL J. Biol. Chem. 264:2118-2125(1989).
RN [2]
RP SEQUENCE OF 226-233; 256-267; 332-341; 667-677 AND 898-908.
RX MEDLINE=86287311; PubMed=3016712;
RA Wels J.J., Fearon D.T., Klickstein L.B., Wong W.W., Richards S.A.,
RA de Bruijn Kops A., Smith J.A., Wels J.H.;
RT "Identification of a partial cDNA clone for the C3d/Epstein-Barr
RT virus receptor of human B lymphocytes: homology with the receptor for
RT fragments C3b and C4b of the third and fourth components of
RT complement.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5639-5643(1986).
RN [3]
RP SEQUENCE OF 492-556 FROM N.A. (SHORT ISOFORM).
RX MEDLINE=93294286; PubMed=8390533;

```





QY 81 LFGEDVFRSAPYVMTVWLRTPACAGRSVY-TE-AYVTIPVGCTCVP-----EPEK 133  
 Db 3035 LGISEESVEGTAVLPGLCKTNPCNGTCTPTTSYV-----CTCAGPYSGDQCEL 3089  
 QY 134 DADSINSIDKQAKLLLPN 154  
 Db 3090 DFDECHSNPCRNATGATVDGFN 3110

RESULT 15  
 CDNL\_HUMAN  
 ID CDNL\_HUMAN STANDARD; PRT; 164 AA.  
 AC P38936;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CYCLIN-DEPENDENT KINASE INHIBITOR 1 (MELANOMA DIFFERENTIATION  
 GN ASSOCIATED PROTEIN 6) (MDA-6) (P21) (CDK-INTERACTING PROTEIN 1).  
 OS CDKN1A OR CDKN1 OR CIP1 OR WAF1 OR MDA6 OR SDI1 OR P21 OR CAP20.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94061996; PubMed=8242751;  
 RA Harper J.W., Adami G.R., Wei N., Keyomarsi K., Elledge S.J.;  
 RT "The p21 Cdk-interacting protein Cipl1 is a potent inhibitor of G1  
 cyclin-dependent kinases.";  
 RL Cell 75:805-816(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94061997; PubMed=8242752;  
 RA El-Deiry W.S., Tokino T., Velculescu V.E., Levy D.B., Parsons R.,  
 RA Trent J.M., Lin D., Mercer W.E., Kinzler K.W., Vogelstein B.;  
 RT "WAF1, a potential mediator of p53 tumor suppression.";  
 RL Cell 75:817-825(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94081955; PubMed=8259214;  
 RA Xiong Y., Hannon G.J., Zhang H., Casso D., Kobayashi R., Beach D.;  
 RT "p21 is a universal inhibitor of cyclin kinases.";  
 RL Nature 366:701-704(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Jiang H., Fisher P.B.;  
 RT "Use of a sensitive and efficient subtraction hybridization protocol  
 for the identification of genes differentially regulated during the  
 induction of differentiation in human melanoma cells.";  
 RL Mol. Cell. Differ. 1:285-299(1993).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Jiang H., Lin J., Herlyn M., Kerbel R.S., Weissman B.E.,  
 RA Welch D.R., Fisher P.B.;  
 RT Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94170884; PubMed=8125163;  
 RA Noda A., Ning Y., Venable S.F., Pereira-Smith O.M., Smith J.R.;  
 RT "Cloning of senescent cell-derived inhibitors of DNA synthesis using  
 an expression screen.";  
 RL Exp. Cell Res. 211:90-98(1994).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95384154; PubMed=7655464;  
 RA Mousses S., Oezcelik H., Lee P.D., Malkin D., Bull S.B.,  
 RA Andrulis I.L.;  
 RT "Two variants of the CIP1/WAF1 gene occur together and are associated  
 with human cancer.";  
 RL Hum. Mol. Genet. 4:1089-1092(1995).  
 RN [8]  
 RP SEQUENCE FROM N.A.

Palmer S.;  
 Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RA RL  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 139-160.  
 RX MEDLINE=97015085; PubMed=8861913;  
 RA Gulbis J.M., Keiman Z., Hurwitz J., O'Donnell M., Kuriyan J.;  
 RT "Structure of the C-terminal region of p21(WAF1/cip1) complexed with  
 human PCNA.";  
 RL Cell 87:297-306(1996).  
 CC -|- FUNCTION: MAY BE THE IMPORTANT INTERMEDIATE BY WHICH P53 MEDIATES  
 ITS ROLE AS AN INHIBITOR OF CELLULAR PROLIFERATION IN RESPONSE TO  
 DNA DAMAGE. MAY BIND TO AND INHIBIT CYCLIN-DEPENDENT KINASE  
 ACTIVITY, PREVENTING PHOSPHORYLATION OF CRITICAL CYCLIN-DEPENDENT  
 KINASE SUBSTRATES AND BLOCKING CELL CYCLE PROGRESSION.  
 CC -|- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -|- TISSUE SPECIFICITY: IS EXPRESSED IN ALL ADULT HUMAN TISSUES,  
 WITH 5-FOLD LOWER LEVELS OBSERVED IN THE BRAIN.  
 CC -|- INDUCTION: BY THE P53 TUMOR SUPPRESSOR/ONCOGENE. ALSO BY MEZEREIN  
 (ANTILEUKEMIC COMPOUND) AND BY IFN-BETA.  
 CC -|- SIMILARITY: THE N-TERMINAL OF CIP1 AND KIP ARE SIMILAR.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; L25610; AAA16109.1; -;  
 DR EMBL; S67388; AAB29246.1; -;  
 DR EMBL; U09579; AAA85641.1; -;  
 DR EMBL; U03106; AAC04313.1; -;  
 DR EMBL; L26165; AAA19811.1; -;  
 DR EMBL; L47233; AAB59560.1; ALT\_INIT.  
 DR EMBL; Z85996; CAB06656.1; -;  
 DR PIR; S39357; S39357.  
 DR SWISS-2DPAGE; P38936; HUMAN.  
 DR MIM; 116899; -;  
 DR InterPro; IPR003175; CDI.  
 DR Pfam; PF02234; CDI; 1.  
 KW Cell cycle; Nuclear protein; Zinc-finger.  
 FT ZN\_FING 13 41  
 FT DOMAIN 141 156 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 SQ SEQUENCE 164 AA; 18119 MW; 98D1E7C519ADFCA9 CRC64;  
 Query Match 8.3%; Score 71.5; DB 1; Length 164;  
 Best Local Similarity 25.2%; Pred. No. 1.7;  
 Matches 39; Conservative 10; Mismatches 53; Indels 53; Gaps 6;  
 QY 30 PAG-----CRPADRRFRPPTNLRVSYPWAYRISYDPRYPVLPAYCLRCGC----- 77  
 Db 4 PAVDRVONPCGSKACRRLLFGVDSQELS-----RDCDALMAGCIOEAR 46  
 QY 78 -----LTGLFGEEDVFRSAP-VYMTVWLRTPACAGRSVYT----- 115  
 Db 47 ERWNFDVFTETPLEGDFAWERVGLGLPKLYLPTGPRGRDELGGRRPTSPALLOTA 106  
 QY 116 -EAYVTIPVGCTCVPPEKQAD-SINSSIDKQGA 148  
 Db 107 EEDHVDLSLSCTLPLVRSGEAEAGSGPGGDSQGRK 141

Search completed: April 14, 2002, 12:41:08  
 Job time: 200 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2002, 12:38:13 ; Search time 24.92 Seconds  
(without alignments)  
939.148 Million cell updates/sec

Title: US-09-731-816-4  
Perfect score: 864  
Sequence: 1 NSARAVLSAFHHTLQLP.....SIDKOGAKLLLPNDAPAGP 160

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL\_17.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhnc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	22.2	197	4 Q9P0M4	Q9p0m4 homo sapien
2	192	22.2	197	4 Q9HC75	Q9hc75 homo sapien
3	184	21.3	178	11 Q9EQ16	Q9eq16 mesocricetu
4	177	20.5	180	11 Q9QXT6	Q9qxt6 mus musculus
5	177	20.5	180	11 Q99WY3	Q99wy3 mus musculus
6	175	20.3	215	11 Q9CTI4	Q9cti4 mus musculus
7	174	20.1	180	4 Q9UHF5	Q9uhf5 homo sapien
8	155	17.9	111	11 Q9EQ17	Q9eq17 rattus norv
9	119.5	13.8	78	4 Q9NUE6	Q9nue6 homo sapien
10	119	13.8	177	4 Q9H293	Q9h293 homo sapien
11	107.5	12.4	151	12 Q40633	Q40633 salmeline
12	99	11.5	148	5 Q19778	Q19778 caenorhabdi
13	91.5	10.6	93	6 Q9BEH5	Q9beh5 equus cabal
14	84.5	9.8	500	4 Q16753	Q16753 homo sapien
15	84.5	9.8	552	4 Q95723	Q95723 homo sapien
16	83.5	9.7	226	5 Q9XUH9	Q9xuh9 caenorhabdi
17	80.5	9.3	805	5 Q9VFD4	Q9vfd4 drosophila
18	80	9.3	944	12 Q9DW96	Q9dw96 rat cytomeg
19	79.5	9.2	520	4 Q9UBG7	Q9ubg7 homo sapien

20	77.5	9.0	1045	6	O46545	O46545 ovis aries
21	77	8.9	354	5	Q9XWJ4	Q9xwj4 caenorhabdi
22	76.5	8.9	733	5	Q9VFX9	Q9vfx9 drosophila
23	76	8.8	242	2	Q9PL24	Q9pl24 chlamydia m
24	76	8.8	1403	12	Q9QP21	Q9qp21 gallid herp
25	75.5	8.7	509	10	O48780	O48780 arabidopsis
26	75	8.7	178	2	Q9X0U1	Q9xou1 thermotoga
27	75	8.7	242	2	O84019	O84019 chlamydia t
28	74.5	8.6	504	2	P72236	P72236 pseudomonas
29	74	8.6	370	11	O08771	O08771 rattus norv
30	74	8.6	496	11	Q9D0H2	Q9doh2 mus musculu
31	74	8.6	1006	5	Q9VBP9	Q9vbp9 drosophila
32	74	8.6	2390	11	O08592	O08592 rattus norv
33	73.5	8.5	195	11	Q9CW28	Q9cw28 mus musculu
34	73.5	8.5	195	12	O87073	O87073 pseudorabie
35	73.3	8.5	577	12	Q9Q3F7	Q9q3f7 pseudorabie
36	73.5	8.5	1087	4	Q14212	Q14212 homo sapien
37	73	8.4	353	10	Q9LHG4	Q9lhg4 arabidopsis
38	73	8.4	903	4	Q9UPX1	Q9upx1 homo sapien
39	72.5	8.4	386	10	Q9ARS6	Q9ars6 oryza sativ
40	72.5	8.4	578	12	Q9PVC0	Q9pvc0 pseudorabie
41	72.5	8.4	2087	11	Q9WUE8	Q9wue8 rattus norv
42	72.5	8.4	2158	11	Q9WU13	Q9wu13 rattus norv
43	72.5	8.4	2167	11	Q9WV48	Q9wv48 rattus norv
44	72	8.3	685	4	O15271	O15271 homo sapien
45	72	8.3	1223	10	Q9SNI5	Q9sni5 oryza sativ

ALIGNMENTS

RESULT 1

ID	Q9P0M4	PRELIMINARY;	PRT;	197 AA.
AC	Q9P0M4;			
DT	01-OCT-2000 (TReMBLrel. 15, Created)			
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)			
DE	INTERLEUKIN 17C.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20105548; Pubmed=10639155;			
RA	Li H., Chen J., Huang A., Stinson J., Heldens S., Foster J., Dowd P.,			
RA	Gurney A.L., Wood W.I.;			
RT	*Cloning and Characterization of IL-17B and IL17C, Two New Members of			
RT	the IL-17 Cytokine Family.*;			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:773-778(2000).			
DR	EMBL: AF152099; AAF28105.1; -			
SQ	SEQUENCE 197 AA; 21765 MW; BAE0152E18DE7D08 CRC64;			

Query Match	22.2%;	Score 192;	DB 4;	Length 197;
Best Local Similarity	34.6%;	Pred. No. 3.8e-12;		
Matches	45;	Conservative 17;	Mismatches 58;	Indels 10; Gaps 4;
QY	7	AVLSAFHHTLQLPREQ-ARNASCPAGGRPADRRFRPTNLRSVSPWAYRISYPARYPR 65		
Db	65	ALVSSLEAASHRHGHERPSATTQCPV-LRP-EEVLEADTHQRSISPRWRYRVDTDDEYRPQ 122		
QY	66	YLPAYCLGCGTGLGEGEDVRFSPAPVYPTVVLARTPACAGRSVYTEA-----Y 118		
Db	123	KLAFACELCGCIDARTGRTAALNSVRLQLSLVLRRRPCSRDGSGLPTPGAFAPFATFTEF 182		
QY	119	VTIPVGCTCV 128		
Db	183	IHPVGCTCV 192		
RESULT	2			

```
Q9HC75
ID Q9HC75 PRELIMINARY; PRT; 197 AA.
AC Q9HC75
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CYTOKINE CX2 PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang W., He L., Wan T., Yuan Z., Cao X.;
RT "Novel human cytokine CX2 with homology to IL-17.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF142410; AAC27921.1; -.
SQ SEQUENCE 197 AA; 21784 MW; BAFBB49F6314A768 CRC64;

Query Match 22.2%; Score 192; DB 4; Length 197;
Best Local Similarity 34.6%; Pred. No. 3.8e-12;
Matches 45; Conservative 17; Mismatches 58; Indels 10; Gaps 4;

Qy 7 AVLSAFHHTLQLGPREQ-ARNASCPAGGRPADRRFRPTNLRVSYPWAYRISYDPARYPR 65
Db 65 ALVSSLEAASHRHGRHSPATTCVP-LRP-EVLEADTHQRSISPNRYRVDTDDEDRYPQ 122
Qy 66 YLPAYCLCRCLGLGEEEDVFRSPAPVYMPVTVLRTACAGRSVYTEA-----Y 118
Db 123 KLAFAECLRCICDARTGRETALNSVRLQLSLVLRRCPSRDGSLPTPGAFATHEF 182
Qy 119 VTIPVGCTV 128
Db 183 IHVPVGCTV 192

RESULT 3
Q9EQI6 PRELIMINARY; PRT; 178 AA.
ID Q9EQI6
AC Q9EQI6
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE NEURONAL INTERLEUKIN-17 RELATED FACTOR (FRAGMENT).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Moore E.E., Presnell S., Garrigues U., Guilbot A., LeGuern E.,
RA Smith D., Yao L., Whitmore T.E., Gilbert T., Kuestner R.E.;
RT "Identification of a novel IL-17 related factor: Demonstration of
RT neuronal expression and evaluation as a candidate for the chromosome
RT 5q-linked form of Charcot-Marie-Tooth disease.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF218725; AAG44134.1; -.
FT NON_TER 178
SQ SEQUENCE 178 AA; 20071 MW; 03D02D60BEC89A51 CRC64;

Query Match 21.3%; Score 184; DB 11; Length 178;
Best Local Similarity 41.1%; Pred. No. 2.3e-11;
Matches 46; Conservative 10; Mismatches 44; Indels 12; Gaps 3;

Qy 23 QARNASCPAGGRPADRRFRPP-----TNLRVSYPWAYRISYDPARYLPYLCRCGCGC 77
Db 72 QLRNSS-----EPAKRRCEVNLQWLNSKRSLSWPGYSINHDPSRIPADLPPEARCLCLGC 126
Qy 78 LTGLGFEEDVFRSPAPVYMPVTVLRR--TPACAGRSVYTEAYVTIPVGCTC 127
```

```
Db 127 VNPFTMOEDRSMSVSPVFSQVPVRRRLCPPPRPGCRHRRVVMETIAGCTC 178

RESULT 4
Q9QXT6 PRELIMINARY; PRT; 180 AA.
ID Q9QXT6
AC Q9QXT6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CYTOKINE-LIKE PROTEIN ZCYT07 (NEURONAL INTERLEUKIN-17 RELATED FACTOR)
DE (INTERLEUKIN 17B).
GN IL17B OR ZCYT07.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Presnell S., Gilbert T., Whitmore T., Foster D., Hart C., Lehner J.,
RA Martinez T., Hoffman R., O'Hara P.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Moore E.E., Presnell S., Garrigues U., Guilbot A., LeGuern E.,
RA Smith D., Yao L., Whitmore T.E., Gilbert T., Kuestner R.E.;
RT "Identification of a novel IL-17 related factor: Demonstration of
RT neuronal expression and evaluation as a candidate for the chromosome
RT 5q-linked form of Charcot-Marie-Tooth disease.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF184570; AAF01319.1; -.
DR EMBL: AF218726; AAG44135.1; -.
DR EMBL: BC002271; AAH02271.1; -.
DR MGD; MGI:1928397; I117b.
SQ SEQUENCE 180 AA; 20309 MW; E26F4C72001997C5 CRC64;

Query Match 20.5%; Score 177; DB 11; Length 180;
Best Local Similarity 40.2%; Pred. No. 1.2e-10;
Matches 45; Conservative 11; Mismatches 46; Indels 10; Gaps 2;

Qy 23 QARNASCPAGGRPADRRFRPTNLRVSYPWAYRISYDPARYLPYLCRCGCLGTLGF 82
Db 72 QLRNSSEPARKKCEVNLQWLNSKRSLSWPGYSINHDPSRIPADLPPEARCLCLGVNPF 131
Qy 83 QEDVFRFRSAPVYMPVTVLRR-----TPACAGRSVYTEAYVTIPVGCTCV 128
Db 132 MQEDRSMSVSPVFSQVPVRRRLCPPPRPGCRQRVVME----TIAVGCTCI 179

RESULT 5
Q99MY3 PRELIMINARY; PRT; 180 AA.
ID Q99MY3
AC Q99MY3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CYTOKINE CX1 PRECURSOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BALB/C;
RA Zhang W., Cao X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF250308; AAK37427.1; -.
SQ SEQUENCE 180 AA; 20268 MW; AAE4CF5FACA3D1LE CRC64;
```

Q9UHF5	PRELIMINARY;	PRT;	180 AA.
Q9UHF5;			
AC	01-MAY-2000 (TReMBLrel. 13, Created)		
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)		
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)		
DT	01-MAR-2001 (TReMBLrel. 16, Last annotation update)		
DE	CYTOKINE-LIKE PROTEIN ZCYT07 (INTERLEUKIN 20) (INTERLEUKIN 17B) (NEURONAL INTERLEUKIN-17 RELATED FACTOR) (INTERLEUKIN-17 BETA).		
DE	ZCYT07 OR IL20.		
GN	Homo sapiens (Human).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Presnell S., Gilbert T., Whitmore T., Foster D., Hart C., Lehner J.,		
RA	Martinez T., Hoffman R., O'Hara P.;		
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20317118; PubMed=10749887;		
RA	Shi Y., Ullrich S.J., Zhang J., Connolly K., Grzegorzewski K.J.,		
RA	Barber M.C., Wang W., Wathen K., Hodge V., Fisher C.L., Olsen H.,		
RA	Ruben S.M., Knyazev I., Cho Y.H., Kao V., Wilkinson K.A.,		
RA	Carrell J.A., Ebner R.;		
RT	"A novel cytokine receptor-ligand pair. Identification, molecular		
RT	characterization, and in vivo immunomodulatory activity.";		
RL	J. Biol. Chem. 275:19167-19176(2000).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20105548; PubMed=10639155;		
RA	Li H., Chen J., Huang A., Stinson J., Heldens S., Foster J., Dowd P.,		
RA	Gurney A.L., Wood W.I.;		
RT	"Cloning and Characterization of IL-17B and IL17C, Two New Members of		
RT	the IL-17 Cytokine Family.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 97:773-778(2000).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RA	Moore E.E., Presnell S., Garrigues U., Guilbot A., LeGuern E.,		
RA	Smith D., Yao L., Whitmore T.E., Gilbert T., Kuestner R.E.;		
RT	"Identification of a novel IL-17 related factor: Demonstration of		
RT	neutrophil expression and evaluation as a candidate for the chromosome		
RT	5q-linked form of Charcot-Marie-Tooth disease.";		
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RA	Zhang W., Wang J., Cao X.;		
RT	"Novel cytokine homology with interleukin-17.";		
RT	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF184969; AAF01318.1; -		
DR	EMBL; AF212311; AAF78775.1; -		
DR	EMBL; AF152098; AAF28104.1; -		
DR	EMBL; AF218727; AAG44136.1; -		
DR	EMBL; AF110385; AAG39637.1; -		
SQ	SEQUENCE 180 AA; 20437 MW; F1B0BC1446D0B14A CRC64;		
Query Match	20.1%;	Score 174;	DB 4; Length 180;
Best Local Similarity	43.7%;	Pred. No. 2.5e-10;	
Matches	38; Conservative	9; Mismatches	38; Indels 2; Gaps
QY	44 TNLRSVSPWAYRISYDPARYPRYLPEAYCLRCGLTGLGEEDVFRFSAPVYMTVLLR 103		
Db	93 SNKRSLSPWGYSINHDPSRIPVDLPEARCLCLGCVNPTMQEDRSMSVVPFSQVPPVRR 152		
QY	104 --TPACAGRSVYTYAYTIPVGTCTV 128		
Db	153 LCPPPPRTGPCRQAVMETIAGCTCI 179		
RESULT	8		
Q9EQ17	PRELIMINARY;	PRT;	111 AA.
AC	Q9EQ17;		
Q9EQ17;			
Query Match	20.3%;	Score 175;	DB 11; Length 215;
Best Local Similarity	40.2%;	Pred. No. 2.4e-10;	
Matches	45; Conservative	10; Mismatches	47; Indels 10; Gaps
QY	23 QARNASCPAGGRPADRRFRPTNLRSVSPWAYRISYDPARYPRYLPEAYCLRCGLTGLF 82		
Db	72 QLRNISEPAKKKEVNLQWLSNKRSLSPWGSYINHDPSRIPADLPEARCLCLGCVNPT 131		
QY	83 GEEDVFRFSAPVYMTVLLR-----TPACAGRSVYTYAYTIPVGTCTV 128		
Db	132 MQEDRSMSVVPFSQVPPVRRRLCPQPPRCQQRVYME-----TIAVGCTCI 179		
RESULT	6		
Q9CT14	PRELIMINARY;	PRT;	215 AA.
ID	Q9CT14		
AC	Q9CT14;		
DT	01-JUN-2001 (TReMBLrel. 17, Created)		
DT	01-JUN-2001 (TReMBLrel. 17, Last sequence update)		
DE	1110006016RIK PROTEIN (FRAGMENT).		
GN	1110006016RIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=EMBRYO;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Karawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner T., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Glassi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,		
RA	Sch		

```

DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE NERF (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Moore E.E., Presnell S., Garrigues U., Guilbot A., LeGuern E.,
RA Smith D., Yao L., Whitmore T.E., Gilbert T., Kuestner R.E.;
RT "Identification of a novel IL-17 related factor: Demonstration of
RT neuronal expression and evaluation as a candidate for the chromosome
RT 5q-linked form of Charcot-Marie-Tooth disease.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF218724; AAG44133.1; -.
FT NON_TER 1 111
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12628 MW; D1598392981BA867 CRC64;

Query Match 17.9%; Score 155; DB 11; Length 111;
Best Local Similarity 43.2%; Pred. No. 1.4e-08;
Matches 35; Conservative 10; Mismatches 36; Indels 0; Gaps 0;

Qy 23 QARNASCPAGGRPADRRFRPTNLSRVSPPAYRISYDPARYPRYLPEAYCLRCGLTGLF 82
Db 9 QLRNSSEPAKKCCVNLQLSNKRSLSIPWGYSINHPSRIPEDLPPEARCLCICGVNPF 68
Qy 83 GEEDVRRSAPVYPTVLR 103
Db 69 MOEDRMSVSPVFSQVPPVRR 89

RESULT 9
Q9NUE6 PRELIMINARY; PRT; 78 AA.
ID Q9NUE6
AC Q9NUE6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE DJ108C2.3 (PUTATIVE NOVEL PROTEIN SIMILAR TO IL17 (INTERLEUKIN 17
DE (CYTOTOXIC T-LYMPHOCYTE-ASSOCIATED SERINE ESTERASE 6)) (CYTOTOXIC T
DE LYMPHOCYTE-ASSOCIATED ANTIGEN 8, CTLA8)) (FRAGMENT).
GN DJ108C2.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Almeida J.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL034343; CAB75300.1; -.
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 78 AA; 8704 MW; 66C3C34E7ACC2790 CRC64;

Query Match 13.8%; Score 119.5; DB 4; Length 78;
Best Local Similarity 37.8%; Pred. No. 4.3e-05;
Matches 28; Conservative 13; Mismatches 28; Indels 5; Gaps 3;

Qy 56 ISYDPARYPRYLPEAYCLRCGLTGLFGEEDVRRSAPVYPTVLRRTTACAGRSV-Y 114
Db 1 VTDPNRYPSVWQAQCNRLGCIINAQ-GKEDISMNSVPIQQTILVRRKHQ---GCSVSF 56
Qy 115 TEAVYTPVGCTCV 128
Db 57 QLEKVLTVGCTCV 70

RESULT 10

```

```

Q9H293 PRELIMINARY; PRT; 177 AA.
ID Q9H293
AC Q9H293
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE INTERLEUKIN 17E.
GN IL17E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J., Ho W.-H., Maruoka M., Corpuz R.T., Baldwin D.T., Foster J.S.,
RA Goddard A.D., Yansura D.G., Vandlen R.L., Wood W.I., Gurney A.L.;
RT "IL-17E, a Novel Proinflammatory Ligand for the IL-17 Receptor Homolog
RT IL-17Rhl.";
RL J. Biol. Chem. 276:1660-1664(2001).
DR EMBL; AF305200; AAG40848.1; -.
SQ SEQUENCE 177 AA; 20330 MW; 52D895710CD59871 CRC64;

Query Match 13.8%; Score 119; DB 4; Length 177;
Best Local Similarity 30.0%; Pred. No. 0.00012;
Matches 36; Conservative 15; Mismatches 59; Indels 10; Gaps 4;

Qy 15 TLQIGPQEQARNASCPAGGRPADRRFRPTNLSRVSPPAYRISYDPARYPRYLPEAYCLC 74
Db 56 TVPVPPLEPARPNRHPSERASE---DGPLNSRAISPPRYELDRNLPLQDLYHARCLC 112
Qy 75 RGCLTGLFGEE-DVRFERSAPVYPTVLRRTTACAGRSVY----TEAVYTPVGCTCV 128
Db 113 PHCVSLQTGSHMDPRGNSSELLYHNQTVFYRRP-CHGEKGTGKCYCLERRLYRSLACVCV 171

RESULT 11
Q40633 PRELIMINARY; PRT; 151 AA.
ID Q40633
AC Q40633
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE INTERLEUKIN 17.
OS Saimiriine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammapherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-488;
RX MEDLINE=98037620; PubMed=9371569;
RA Knappe A., Hiller C., Thureau M., Wittmann S., Hofmann H.,
RA Fleckenstein B., Fickscher H.;
RT "The superantigen-homologous viral immediate-early gene iel4/vsag in
RT herpesvirus saimiri-transformed human T cells.";
RL J. Virol. 71:9124-9133(1997).
DR EMBL; Y13183; CAA73627.1; -.
SQ SEQUENCE 151 AA; 17189 MW; 49B2C9430C46BE32 CRC64;

Query Match 12.4%; Score 107.5; DB 12; Length 151;
Best Local Similarity 31.4%; Pred. No. 0.0015;
Matches 27; Conservative 12; Mismatches 36; Indels 11; Gaps 3;

Qy 47 RSVSPWAYRISYDPARYPRYLPEAYCLRCGLTGLFGEEDVRRSAPVYPTVLRTPA 106
Db 65 RSTSPWTLRYNEDQDRYPSVWEAKRYLGCVNA-DGNVDYHMNSVPIQOEILVVRK--- 120
Qy 107 CAGRSVYTEAY----VTIPVGCTCV 128
Db 121 ---GHNPCPNPSFRLEKMLTVTGCTCV 143

```

DR	EMBL; AY014959; AAG52985.1; -	NON_TER	1	1	
FT		NON_TER	93	93	
FT		NON_TER	93	93	
SO	SEQUENCE	93 AA;	10635 MW;	CC2DA9F02302129F	CRC64;

RESULT	15
O95723	
ID	O95723 PRELIMINARY; PRT; 552 AA.
AC	O95723;
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	DJ453C12.2 (SIMILAR TO TRANSCRIPTION FACTOR RPB-L).
GN	DJ453C12.2.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI	TAXID=9606;

```
[1]
RN  SEQUENCE FROM N.A.
RP  Bridgeman A.;
RA  Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RL  EMBL; AL021578; CAA16521.1; -.
DR  InterPro; IPR002909; IPT_TIG.
DR  Pfam; PF01833; TIG; 1.
SQ  SEQUENCE 552 AA; 60422 MW; 5E2EC4B12237AF43 CRC64;

Query Match      9.8%; Score 84.5; DB 4; Length 552;
Best Local Similarity 28.1%; Pred. No. 1.5;
Matches 39; Conservative 11; Mismatches 44; Indels 45; Gaps 8;

QY  18 LGPREQARNASCPAGGRPADRRFRPPTNLRVSPWAYRISYDPARYPYRL-----PEAYCL 73
Db   ||| ||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  436 LGPGEQGR-----GVHASSESPQ-----SPWC-----STPRSPRSLVCVVPDVAAF 477
Db   ||| ||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY  74 CRGCLTGLFGEDVFRSAPVYMTPTVLRRTPACAGG-----RSVYTEAYVTIPVGCT 126
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  478 C-----SDRWLRAPITIPMSLVR-----ADGLFYPSAFSFTYTPEYSVRP-GHP 521
Db   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY  127 CVPEPEKDADSINSSIDKQ 145
Db   ||| ||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY  522 GVPEPATDADALLESIHQE 540
```

Search completed: April 14, 2002, 12:41:39  
Job time: 206 sec



```
Db 16 SEAAARKIPKVGHTFFQ-----KPESCPVPVGGSMKLDIGIINENQRYVMSRNIERS 68
QY 49 VSPWAYRISYDPARYPRYLPEAYCLRCGLTGLFGEDVFRSAPVYMPVTWVLRRTPACA 108
   ||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db 69 TSPWNVTWTDNRPYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQOETLVVRRKHQ-- 125
QY 109 GGRSV-YTEAYVTIPVGCTCV 128
   ||| |::||| |::||| |::||| |::||| |::||| |::|||
Db 126 -GCSVSFQLEKVLVTGCTCV 145

RESULT 2
US-08-833-823-12
; Sequence 12, Application US/08033823
; Patent No. 5969093
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Kelleher, Kerry
; APPLICANT: Carlin, McKeough
; TITLE OF INVENTION: DNA SEQUENCES AND SECRETED PROTEINS
; TITLE OF INVENTION: ENCODED THEREBY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc. -- Legal Affairs
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 10-APR-1997
; APPLICATION NUMBER: US/08/833,823
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/514,014
; FILING DATE: 11-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: GI6000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-823-12

Query Match 18.4%; Score 159; DB 2; Length 153;
Best Local Similarity 32.6%; Pred. No. 7.9e-11;
Matches 46; Conservative 18; Mismatches 53; Indels 24; Gaps 7;

QY 1 NSARARAVLSAFHHTLQLGPREQARNASC---PAGGRPAD-----RRFRPTNL--RS 48
   :||| |::||| |::||| |::||| |::||| |::||| |::|||
Db 16 SEAAARKIPKVGHTFFQ-----KPESCPVPVGGSMKLDIGIINENQRYVMSRNIERS 68
QY 49 VSPWAYRISYDPARYPRYLPEAYCLRCGLTGLFGEDVFRSAPVYMPVTWVLRRTPACA 108
   ||| |::||| |::||| |::||| |::||| |::||| |::|||
Db 69 TSPWNVTWTDNRPYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQOETLVVRRKHQ-- 125
QY 109 GGRSV-YTEAYVTIPVGCTCV 128
   ||| |::||| |::||| |::||| |::||| |::||| |::|||
Db 126 -GCSVSFQLEKVLVTGCTCV 145

RESULT 4
US-08-685-239-2
; Sequence 2, Application US/08685239
```

```
RESULT 3
US-09-034-810-2
; Sequence 2, Application US/09034810
; Patent No. 6043344
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: Kelleher, Kerry
; APPLICANT: Carlin, McKeough
; APPLICANT: Goldman, Samuel
; APPLICANT: Pittman, Debra
; APPLICANT: Mi, Sha
; APPLICANT: Neben, Steven
; APPLICANT: Giannotti, JoAnn
; APPLICANT: Golden/Fleet, Margaret
; TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,810
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/685,239
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: GI5262
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-034-810-2

Query Match 18.4%; Score 159; DB 3; Length 163;
Best Local Similarity 32.6%; Pred. No. 8.5e-11;
Matches 46; Conservative 18; Mismatches 53; Indels 24; Gaps 7;

QY 1 NSARARAVLSAFHHTLQLGPREQARNASC---PAGGRPAD-----RRFRPTNL--RS 48
   :||| |::||| |::||| |::||| |::||| |::||| |::|||
Db 26 SEAAARKIPKVGHTFFQ-----KPESCPVPVGGSMKLDIGIINENQRYVMSRNIERS 78
QY 49 VSPWAYRISYDPARYPRYLPEAYCLRCGLTGLFGEDVFRSAPVYMPVTWVLRRTPACA 108
   ||| |::||| |::||| |::||| |::||| |::||| |::|||
Db 79 TSPWNVTWTDNRPYPSEVVQAQCRNLGCINAQ-GKEDISMNSVPIQOETLVVRRKHQ-- 135
QY 109 GGRSV-YTEAYVTIPVGCTCV 128
   ||| |::||| |::||| |::||| |::||| |::||| |::|||
Db 136 -GCSVSFQLEKVLVTGCTCV 155

RESULT 4
US-08-685-239-2
; Sequence 2, Application US/08685239
```

Patent No. 6074849  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: Kelleher, Kerry  
APPLICANT: Carlin, McKeough  
APPLICANT: Goldman, Samuel  
APPLICANT: Pittman, Debra  
APPLICANT: Mi, Sha  
APPLICANT: Neben, Steven  
APPLICANT: Giannotti, JoAnn  
APPLICANT: Golden Fleet, Margaret  
TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/685,239  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: G15262  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 163 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-685-239-2

Query Match 18.48; Score 159; DB 3; Length 163;  
Best Local Similarity 32.68; Pred. No. 8.5e-11;  
Matches 46; Conservative 18; Mismatches 53; Indels 24; Gaps 7;  
Qy 1 NSARAVLSAFHHTLQLGPREQARNASC---PAGGRPAD-----RRFRPTNL--RS 48  
Db 26 SEAAARKTPKVCHTFQ-----KPESCVPVPGSMKLDGIINENORVSMRSNIESRS 78  
Qy 49 VSPWAYRISDPARYPRYLPCALCGLGFGEEEDVFRSAPVYMTVVLRRTPACA 108  
Db 79 TSPWNYTVTDPNRYPSEVQAOGRNLGINAQ-GKEDISMNSVPIQOETLWVRKHQ-- 135  
Qy 109 GGRSV-YTEAVTTPVGCTCV 128  
Db 136 -GCSVSFQLEKVLTVGCTCV 155

RESULT 5  
US-08-432-994A-8  
Sequence 8, Application US/08432994A  
Patent No. 6274711  
GENERAL INFORMATION:  
APPLICANT: Golstein, Pierre  
APPLICANT: Rouvier, Eric  
APPLICANT: Fossiez, Francois  
APPLICANT: Lebecque, Serge J.E.  
APPLICANT: Djossou, Odile  
APPLICANT: Bancheureau, Jacques

TITLE OF INVENTION: PURIFIED MAMMALIAN CTLA-8 ANTIGENS AND  
TITLE OF INVENTION: RELATED REAGENTS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/432,994A  
FILING DATE: 02-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/250,846  
FILING DATE: 27-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/177,747  
FILING DATE: 05-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,203  
FILING DATE: 14-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0388K3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-852-9196  
TELEFAX: 415-496-1200  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 155 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-432-994A-8

Query Match 15.78; Score 135.5; DB 4; Length 155;  
Best Local Similarity 32.38; Pred. No. 4e-08;  
Matches 41; Conservative 9; Mismatches 46; Indels 31; Gaps 7;  
Qy 25 RNACSPAGGRPADRRFRPT---NL-----RSVSPWAYRISYDPAR 62  
Db 29 RNPCCP---NSEDKNF-PTVMVNLNHNRTNTPKRSSDYNNRSTSPWNLRRNEDPER 84  
Qy 63 YPRYLPCAYLCRGLTGLFGEEEDVFRSAPVYMTVVLRRTPA-CAGRSVYTBAYVTI 121  
Db 85 YPSVWEAKRHLGICINA-DGNVDYHMNSVPVQIQEILVLRREPPHCPNS---FRLEKILV 140  
Qy 122 PVGCTCV 128  
Db 141 SVGCTCV 147

RESULT 6  
US-08-432-994A-10  
Sequence 10, Application US/08432994A  
Patent No. 6274711  
GENERAL INFORMATION:  
APPLICANT: Golstein, Pierre  
APPLICANT: Rouvier, Eric  
APPLICANT: Fossiez, Francois  
APPLICANT: Lebecque, Serge J.E.  
APPLICANT: Djossou, Odile  
APPLICANT: Bancheureau, Jacques  
TITLE OF INVENTION: PURIFIED MAMMALIAN CTLA-8 ANTIGENS AND

;; TITLE OF INVENTION: RELATED REAGENTS  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: DNAX Research Institute  
;; STREET: 901 California Avenue  
;; CITY: Palo Alto  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94304-1104  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/432,994A  
;; FILING DATE: 02-MAY-1995  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/250,846  
;; FILING DATE: 27-MAY-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/177,747  
;; FILING DATE: 05-JAN-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/077,203  
;; FILING DATE: 14-JUN-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Ching, Edwin P.  
;; REGISTRATION NUMBER: 34,090  
;; REFERENCE/DOCKET NUMBER: DX0388K3  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-852-9196  
;; TELEFAX: 415-496-1200  
;; INFORMATION FOR SEQ ID NO: 10:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 158 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-432-994A-10

Query Match 13.1%; Score 113.5; DB 4; Length 158;  
Best Local Similarity 31.4%; Pred. No. 1.4e-05;  
Matches 37; Conservative 16; Mismatches 46; Indels 19; Gaps 6;  
QY 32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLRCGCLTGLFGEDVRF 90  
| : : | | | | | | | | | | | | : : | : |  
Db 58 GAKVSSR--RPSDYLNRSTSPWTLHRNEDPDYPSVWEAQRHRCVNA-EGKLDHNM 114  
| : : | | | | | : : | | | | : | : | : | : |  
QY 91 SAPVYMTVTLRRTP-ACAGRSVYTEAYVTPVGCTVPEKADSISSDKOGA 147  
| : : | | | | | : : | | | | | : | : | : | : |  
Db 115 SVLIQQLVLRKPEPSCP---FTFRVEKMLVGVGCTCV-----ASIVRQA 158

RESULT 7  
US-08-620-694A-7  
; Sequence 7, Application US/08620694A  
; Patent No. 5869286  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: Apple Power Macintosh  
;; OPERATING SYSTEM: Apple Operating System 7.5.5  
;; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/620,694A  
;; FILING DATE: 21 MARCH 1996  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: USSN 08/538,765  
;; FILING DATE: 7 AUGUST 1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: USSN 08/410,535  
;; FILING DATE: 23 MARCH 1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Perkins, Patricia Anne  
;; REGISTRATION NUMBER: 34,695  
;; REFERENCE/DOCKET NUMBER: 2617-B  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206)587-0430  
;; TELEFAX: (206)  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 158 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-620-694A-7

Query Match 13.0%; Score 112; DB 2; Length 158;  
Best Local Similarity 33.3%; Pred. No. 2.1e-05;  
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;  
QY 32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPARYPRYLPEAYCLRCGCLTGLFGEDVRF 90  
| : : | | | | | | | | | | | | : : | : |  
Db 58 GAKVSSR--RPSDYLNRSTSPWTLHRNEDPDYPSVWEAQRHRCVNA-EGKLDHNM 114  
| : : | | | | | : : | | | | : | : | : | : |  
QY 91 SAPVYMTVTLRRTP-ACAGRSVYTEAYVTPVGCTCV 128  
| : : | | | | | : : | | | | | : | : | : | : |  
Db 115 SVLIQQLVLRKPEPSCP---FTFRVEKMLVGVGCTCV 150

RESULT 8  
US-09-022-255-7  
; Sequence 7, Application US/09022255  
; Patent No. 6072033  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/022,255  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:



```

;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-09-022-253-7

```

Query Match 13.0%; Score 112; DB 3; Length 158;  
Best Local Similarity 33.3%; Pred. No. 2.1e-05;  
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;

Qy	32	GRPADRRFRPPNLT-RSVSPWAYRTSYDDARYPRYLPEAYCLCRGCLTGLFGFEEDVRFR	90
		: : : :   : : : :   : : : :   : : : :   : : : :   : : : :   : : : :	
Db	58	GAKVSSR--RPSDYLNKSTSPWTLHRNEDPDYPSVIWEAQCRHQRCVNA-EGKLDHHMN	114
Qy	91	SAPYMFVTVYLRRTPT-ACAGGRSVYTEAYVTIPVGCTCV	128
Db	115	SVLTIOBILVLKREPESCP---ETFERVEKMLVGVGCTCV	150

RESULT 11  
US-09-022-260-7  
; Sequence 7, Application US/09022260  
; Patent No. 6100235  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,260
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995

```

ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-260-7

Query Match 13.0%; Score 112; DB 3; Length 158;  
Best Local Similarity 33.3%; Pred. No. 2.1e-05;  
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;

QV 32 GGRPADRRFRPPTNL-RSVSPWAYRISYDPPARYPRYLPEAYCLCRGCLTGLFGCEEDVFR 90

Db 58 GAKVSSR--RPSDYLNRESTSPWTLHRNEDPRPVSIVTEAQRHORCVNA-EGKLDHNM 114  
QY 91 SAPYMYMTVLRLRTP-ACAGRSVYTEAYVIPVGTCV 128  
Db 115 SVLTQOEILLVKRPESC---FTFRKMLVGVGTCV 150

```

RESULT 12
US-09-022-359-7
; Sequence 7, Application US/09022259
; Patent No. 6191104
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhenqin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6191104el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
;

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,259
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995

```

ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

Query Match 13.0%; Score 112; DB 4; Length 158;  
Best Local Similarity 33.3%; Pred. NO. 2.1e-05;  
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;

[illegible]

RESULT 13  
US-09-022-257-7

```
; Sequence 7, Application US/09022257
; Patent No. 6197525
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fauslow, William
; TITLE OF INVENTION: No. 6197525el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,257
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/620,694
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-022-257-7

Query Match 13.0%; Score 112; DB 4; Length 158;
Best Local Similarity 33.3%; Pred. No. 2.1e-05;
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;

QY 32 GGRPADRRFRPPNRL-RSVSPWAYRISYDPARYPRYLPEAYCLRCGLTGLFGDEEDVRF 90
| : : | | | | | : : | | | | | : : | | : |
Db 58 GAKVSSR--RPSDYNKRSPTWTLHRNEDPDYPSVWEAKQCRHQRVCNA-EGKLDHMH 114
| : : | | | | | : : | | | | | : : | | : |

QY 91 SAPVYMTVVLRRTP-CAGGRSVYTEAYVTIPVGCTCV 128
| : : | | | | | : : | | | | | : : | | : |
Db 115 SVLIQEIILVKREPESCP---FTFRVEKMLVGVGCTCV 150
| : : | | | | | : : | | | | | : : | | : |

RESULT . 14
US-08-620-694A-8
; Sequence 8, Application US/08620694A
; Patent No. 5869286
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fauslow, William
; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
```

```
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,694A
; FILING DATE: 21 MARCH 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Herpesvirus Saimiri
; STRAIN: ORF13
; US-08-620-694A-8

Query Match 12.6%; Score 108.5; DB 2; Length 151;
Best Local Similarity 31.4%; Pred. No. 5e-05;
Matches 27; Conservative 11; Mismatches 37; Indels 11; Gaps 3;

QY 47 RSVSPWAYRISYDPARYPRYLPEAYCLRCGLTGLFGDEEDVRFSAVPYMTVVLRRTPA 106
| | | | | : | | | | | : | | | | | : | | : |
Db 65 RSTSPWTLHRNEDQDRYPSVWEAKCRYLGCVNA-DGNVDYHMNSVPITQQEILVVRK--- 120
| | | | | : | | | | | : | | | | | : | | : |

QY 107 CAGGRSVYTEAY----VTIPVGCTCV 128
| : : | | | | | : : | | | | | : : | | : |
Db 121 ---GHQPCPNFSRLEKMLVTGVGCTCV 143
| : : | | | | | : : | | | | | : : | | : |

RESULT 15
US-09-034-810-6
; Sequence 6, Application US/09034810
; Patent No. 6043344
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: Kelleher, Kerry
; APPLICANT: Carlin, McKeough
; APPLICANT: Goldman, Samuel
; APPLICANT: Pittman, Debra
; APPLICANT: Mi, Sha
; APPLICANT: Neben, Steven
; APPLICANT: Giannotti, JoAnn
; APPLICANT: Golden Fleet, Margaret
; TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins
; NUMBER OF SEQUENCES: 9
```

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Genetics Institute, Inc.  
;; STREET: 87 CambridgePark Drive  
;; CITY: Cambridge  
;; STATE: Massachusetts  
;; COUNTRY: USA  
;; ZIP: 02140  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/034,810  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/685,239  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Brown, Scott A.  
;; REGISTRATION NUMBER: 32,724  
;; REFERENCE/DOCKET NUMBER: G15262  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 498-8224  
;; TELEFAX: (617) 876-5851  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 151 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-09-034-810-6

Query Match 12.6%; Score 108.5; DB 3; Length 151;  
Best Local Similarity 31.4%; Pred. No. 5e-05;  
Matches 27; Conservative 11; Mismatches 37; Indels 11; Gaps 3;  
Qy 47 RSVSPWAYRISYDPARYPRYLPEAYCLRCGLTCLGCEEDVRFSPAPVYMTVVLRTTPA 106  
Db 65 RSTSPWTLHRNEDODRPSVWEAKCRYLGCVNA-DGNVDYHMNSVPQQEILVVRK--- 120  
Qy 107 CAGGRSVYTEAY----VTIPVGCTCV 128  
Db 121 ---CHQPCPNSEFRLEKMLVTVGCTCV 143

Search completed: April 14, 2002, 12:37:45  
Job time: 22 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2002, 12:38:38 ; Search time 23.92 Seconds  
(without alignments)  
495.474 Million cell updates/sec

Title: US-09-731-816-4

Perfect score: 160

Sequence: 1 NSARARVLSAFHTLQLGP.....SIDKQAKLLGPNDAPAGP 160

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 6

Total number of hits satisfying chosen parameters: 684

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A\_Geneseq\_1101.\*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160	100.0	160	21	AAV53891
2	153	95.6	173	21	AAV53893
3	153	95.6	185	21	AAV70656
4	153	95.6	186	21	AAV70655
5	153	95.6	187	21	AAV70654
6	153	95.6	187	21	AAV70663
7	153	95.6	202	21	AAV70695
8	153	95.6	202	21	AAV70689
9	153	95.6	202	21	AAV70653
10	153	95.6	209	21	AAV70662
11	124	77.5	202	22	AAU04953

12	97	60.6	151	21	AAV70594
13	97	60.6	151	21	AAV70690
14	57	35.6	57	21	AAV70666
15	54	33.8	54	21	AAV70664
16	49	30.6	103	21	AAV70888
17	34	21.2	34	21	AAV70861
18	27	16.9	183	21	AAV70658
19	27	16.9	205	21	AAV70597
20	27	16.9	205	21	AAV70691
21	27	16.9	205	21	AAV70657
22	27	16.9	205	22	AAV5341
23	23	14.4	31	21	AAV70659
24	24	12.5	21	21	AAV70660
25	18	11.2	44	21	AAV70596
26	16	10.0	16	21	AAV70665
27	14	8.8	14	21	AAV70667
28	8	5.0	805	21	AAV84562
29	8	5.0	805	21	AAV67310
30	8	5.0	805	22	AAV72667
31	8	5.0	805	22	AAV84095
32	7	4.4	47	18	AAV17971
33	7	4.4	51	21	AAV32429
34	7	4.4	72	21	AAV02095
35	7	4.4	74	20	AAV14475
36	7	4.4	84	21	AAV53775
37	7	4.4	87	21	AAV53890
38	7	4.4	90	21	AAV65590
39	7	4.4	114	21	AAV52552
40	7	4.4	123	21	AAV07601
41	7	4.4	123	21	AAV07683
42	7	4.4	170	19	AAV98374
43	7	4.4	197	21	AAV18911
44	7	4.4	197	21	AAV07602
45	7	4.4	197	21	AAV07684
46	7	4.4	197	21	AAV22338
47	7	4.4	197	21	AAV44460
48	7	4.4	197	21	AAV53892
49	7	4.4	197	22	AAU04951
50	7	4.4	206	21	AAV44485
51	7	4.4	261	20	AAV38934
52	7	4.4	307	18	AAV20285
53	7	4.4	307	18	AAV24626
54	7	4.4	316	18	AAV20864
55	7	4.4	346	22	AAV82710
56	7	4.4	367	21	AAV58112
57	7	4.4	425	21	AAV44462
58	7	4.4	426	22	AAV16773
59	7	4.4	426	22	AAV04490
60	7	4.4	467	21	AAV95647
61	7	4.4	467	21	AAV75495
62	7	4.4	467	21	AAV75496
63	7	4.4	468	21	AAV75497
64	7	4.4	582	19	AAV52134
65	7	4.4	712	17	AAV95648
66	7	4.4	1580	18	AAV23719
67	7	4.4	1580	18	AAV22609
68	6	3.8	13	6	AAV50642
69	6	3.8	24	20	AAV55900
70	6	3.8	26	20	AAV36616
71	6	3.8	26	21	AAV09177
72	6	3.8	34	22	AAV59486
73	6	3.8	41	21	AAV58882
74	6	3.8	47	22	AAV13958
75	6	3.8	47	22	AAV26364
76	6	3.8	47	22	AAV01701
77	6	3.8	50	22	AAV19041
78	6	3.8	50	22	AAV31679
79	6	3.8	60	22	AAV31176
80	6	3.8	61	22	AAV14221
81	6	3.8	64	21	AAV59891
82	6	3.8	75	21	AAV24523
83	6	3.8	75	21	AAV50401
84	6	3.8	75	21	AAV50421

A human interleukin  
A rat interleukin-  
Human transforming  
Human transforming  
A human interleukin  
Human transforming  
Mature murine tran  
A murine interleuk  
A murine interleuk  
Murine transformin  
Mouse 2CYT07 prote  
Human transforming  
Human transforming  
A rat interleukin  
Human transforming  
Human transforming  
A human angiotensi  
Human MPR015 amin  
Human angiotensin  
Human zace2 protel  
RAC protein kinase  
Human secreted pro  
Cyclohexanone mono  
Fragment of human  
Human colon cancer  
Partial amino acid  
Human prostate can  
Helicobacter pylor  
A human interleuk  
A human interleuk  
H. pylori GHPO 112  
A novel polypeptid  
A human interleuk  
A human interleuk  
Human interleukin-  
Human interleukin  
Amino acid sequenc  
Human Interleukin  
Human Interleukin  
Neisseria gonorrhoe  
H. pylori cytoplas  
H. pylori cytoplas  
H. pylori cytoplas  
S. epidermidis ope  
Haemorrhagic enter  
Human Interleukin  
Peptide #3207 enco  
Peptide #3172 enco  
Neisseria meningit  
Neisseria gonorrhoe  
Neisseria meningit  
Neisseria meningit  
Rabbit membrane-ty  
Thermostable DNA-1  
Platenolide syntha  
Platenolide syntha  
Sequence (b) of a  
Oligopeptide from  
Fragment of human  
Hepatitis GB virus  
Human secreted pro  
Human secreted pro  
Polyketide polylin  
Peptide #392 enco  
Peptide #401 enco  
Peptide #383 enco  
Peptide #5475 enco  
Peptide #5716 enco  
Peptide #5213 enco  
Human novel protei  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia

85	6	3.8	76	21	AAG34320	Arabidopsis thalia	158	6	3.8	163	22	AAG82618	S. epidermidis ope
86	6	3.8	78	22	AAAM16070	Peptide #2504 enco	159	6	3.8	163	22	AAAB87605	Human PRO20110. H
87	6	3.8	78	22	AAAM28564	Peptide #2601 enco	160	6	3.8	163	22	AAAB42209	Human OREF ORF1973
88	6	3.8	78	22	AAAM03797	Peptide #2479 enco	161	6	3.8	167	21	AAAG18653	Arabidopsis thalia
89	6	3.8	79	16	AAAR76572	Human CTIA-8 exon-	162	6	3.8	167	21	AAAG18652	Arabidopsis thalia
90	6	3.8	80	21	AAAG34319	Sequence encoded b	163	6	3.8	168	21	AAAB54060	Human pancreatic c
91	6	3.8	83	6	AAAF50649	Sequence encoded b	164	6	3.8	169	21	AAAG18652	Human pancreatic c
92	6	3.8	83	21	AAAG24522	Arabidopsis thalia	165	6	3.8	169	21	AAAG18652	Human pancreatic c
93	6	3.8	83	21	AAAG50400	Arabidopsis thalia	166	6	3.8	177	16	AAAR75479	Hepatitis GB virus
94	6	3.8	83	21	AAAG50420	Arabidopsis thalia	167	6	3.8	177	16	AAAR75479	Hepatitis GB virus
95	6	3.8	83	21	AAAG50400	Arabidopsis thalia	168	6	3.8	177	22	AAAB63279	Hepatitis GB virus
96	6	3.8	86	22	AAAY58037	Human EST encoded	169	6	3.8	178	22	AAAB63279	Human breast cance
97	6	3.8	88	22	AAAM23608	Human EST encoded	170	6	3.8	178	22	AAAB63448	Human breast cance
98	6	3.8	94	9	AAAP80233	Sequence of the fi	171	6	3.8	179	20	AAAY39342	Staphylococcus aur
99	6	3.8	100	21	AAAB22788	Caulobacter sp. FW	172	6	3.8	179	20	AAAY39342	S. aureus spoJ2 p
100	6	3.8	100	21	AAAG03539	Human secreted pro	173	6	3.8	181	22	AAAM13581	Peptide #15 enco
101	6	3.8	104	22	AAAG75670	Human colon cancer	174	6	3.8	181	22	AAAM13581	Peptide #15 enco
102	6	3.8	105	20	AAAB87706	A cysteine rich so	175	6	3.8	181	22	AAAM25979	Peptide #16 enco
103	6	3.8	105	21	AAAY32329	PRO1081 polypeptid	176	6	3.8	181	22	AAAM25979	Peptide #16 enco
104	6	3.8	105	22	AAAB47291	Arabidopsis thalia	177	6	3.8	182	20	AAAY52783	Treponema pallidum
105	6	3.8	109	21	AAAG22176	Arabidopsis thalia	178	6	3.8	187	21	AAAG35441	zea mays protein f
106	6	3.8	109	21	AAAG34643	Arabidopsis thalia	179	6	3.8	187	21	AAAG35441	zea mays protein f
107	6	3.8	109	21	AAAY32327	His(8)-mouse F1221	180	6	3.8	190	21	AAAG35440	Arabidopsis thalia
108	6	3.8	110	21	AAAG34642	Arabidopsis thalia	181	6	3.8	190	21	AAAG35440	Arabidopsis thalia
109	6	3.8	111	20	AAAB87704	A cysteine rich so	182	6	3.8	205	21	AAAG59875	Arabidopsis thalia
110	6	3.8	111	21	AAAB42801	Human OREF ORF2565	183	6	3.8	205	21	AAAG59875	Arabidopsis thalia
111	6	3.8	111	21	AAAB33464	Human PRO720 prote	184	6	3.8	218	17	AAAG92223	Paracoccus denitri
112	6	3.8	111	21	AAAY32328	Mouse F1221 inhibi	185	6	3.8	221	21	AAAG17627	Arabidopsis thalia
113	6	3.8	111	22	AAAG81818	S. epidermidis ope	186	6	3.8	221	21	AAAG17627	Arabidopsis thalia
114	6	3.8	119	19	AAAY21424	Human high mobil	187	6	3.8	226	21	AAAY44938	Artichoke sulphate
115	6	3.8	122	21	AAAG22175	Arabidopsis thalia	188	6	3.8	229	21	AAAB11637	A. vitis hypersens
116	6	3.8	125	21	AAAG26758	zea mays protein f	189	6	3.8	235	21	AAAB32504	S. lavendulae Mit
117	6	3.8	130	21	AAAB21174	Mouse partial IL-3	190	6	3.8	237	21	AAAG23623	Arabidopsis thalia
118	6	3.8	132	16	AAAR76574	Human CTIA-8 matur	191	6	3.8	237	21	AAAG23623	Arabidopsis thalia
119	6	3.8	142	22	AAAB78956	C. glutamicum SRT	192	6	3.8	244	20	AAAR26789	Human 7TM clone HD
120	6	3.8	146	20	AAAB36615	Fragment of human	193	6	3.8	246	20	AAAR26789	Human 7TM clone HD
121	6	3.8	148	21	AAAB41467	Human OREF ORF1231	194	6	3.8	246	20	AAAR26789	Human 7TM clone HD
122	6	3.8	148	21	AAAB41467	Human OREF ORF1231	195	6	3.8	253	21	AAAG10426	Arabidopsis thalia
123	6	3.8	150	16	AAAR76570	Mouse CTIA-8. Mus	196	6	3.8	256	22	AAAB63478	Human breast cance
124	6	3.8	150	18	AAAM13652	Rat CTIA-8. Rattu	197	6	3.8	256	22	AAAB63478	Human breast cance
125	6	3.8	150	21	AAAB03836	Rat cytotoxic T ly	198	6	3.8	259	20	AAAY38931	Neisseria meningit
126	6	3.8	150	21	AAAB03836	Rat CTIA-8 protein	199	6	3.8	259	20	AAAY38931	Neisseria meningit
127	6	3.8	151	16	AAAR76571	Herpesvirus ORF13	200	6	3.8	264	20	AAAY43918	Drosophila protein
128	6	3.8	151	17	AAAM02387	HVS13 (viral homol	201	6	3.8	269	16	AAAR81994	Variant lipase of
129	6	3.8	151	18	AAAM13653	Herpesvirus Saimir	202	6	3.8	269	16	AAAR81994	Variant lipase of
130	6	3.8	151	20	AAAM92413	Herpesvirus Saimir	203	6	3.8	269	16	AAAR81996	Variant lipase of
131	6	3.8	151	21	AAAY99940	Herpesvirus Saimir	204	6	3.8	269	16	AAAR81997	Variant lipase of
132	6	3.8	151	21	AAAY97134	Herpesvirus Saimir	205	6	3.8	269	16	AAAR81999	Variant lipase of
133	6	3.8	151	21	AAAY97134	Herpesvirus Saimir	206	6	3.8	269	16	AAAR82000	Variant lipase of
134	6	3.8	151	21	AAAY97185	Herpesvirus Saimir	207	6	3.8	269	16	AAAR82001	Variant lipase of
135	6	3.8	151	21	AAAB03811	Fragment of HVS13	208	6	3.8	269	16	AAAR82001	Variant lipase of
136	6	3.8	151	21	AAAB03837	Herpesvirus Saimir	209	6	3.8	269	16	AAAR81992	Variant lipase of
137	6	3.8	151	22	AAAM25492	Human protein sequ	210	6	3.8	269	16	AAAR81992	Variant lipase of
138	6	3.8	151	22	AAAB62065	HSV ORF13 polypept	211	6	3.8	269	16	AAAR82011	Variant lipase of
139	6	3.8	151	22	AAAY72753	Herpesvirus Saimir	212	6	3.8	269	16	AAAR82012	Variant lipase of
140	6	3.8	153	18	AAAM28514	Product of clone B	213	6	3.8	269	16	AAAR82013	Variant lipase of
141	6	3.8	153	21	AAAB41569	Human OREF ORF1333	214	6	3.8	269	16	AAAR82014	Variant lipase of
142	6	3.8	155	16	AAAR76573	Human CTIA-8. Hom	215	6	3.8	269	16	AAAR82015	Variant lipase of
143	6	3.8	155	22	AAAB83939	Human secreted pro	216	6	3.8	269	16	AAAR82016	Variant lipase of
144	6	3.8	155	22	AAAU00890	Human cancer relat	217	6	3.8	269	16	AAAR82017	Variant lipase of
145	6	3.8	158	17	AAAM02386	Murine CTIA8 (inte	218	6	3.8	269	16	AAAR82003	Variant lipase of
146	6	3.8	158	20	AAAY92412	Murine CTIA-8 prot	219	6	3.8	269	16	AAAR82004	Variant lipase of
147	6	3.8	158	21	AAAY99939	Murine CTIA-8 prot	220	6	3.8	269	16	AAAR82005	Variant lipase of
148	6	3.8	158	21	AAAY97135	Murine interleukin	221	6	3.8	269	16	AAAR82006	Variant lipase of
149	6	3.8	158	21	AAAY97184	Murine interleukin	222	6	3.8	269	16	AAAR82007	Variant lipase of
150	6	3.8	158	21	AAAB03810	Fragment of murine	223	6	3.8	269	16	AAAR82008	Variant lipase of
151	6	3.8	158	22	AAAB62064	Murine CTIA-8 poly	224	6	3.8	269	16	AAAR82009	Variant lipase of
152	6	3.8	158	22	AAAY72752	Murine Interleukin	225	6	3.8	269	16	AAAR82010	Variant lipase of
153	6	3.8	162	22	AAAG81760	S. epidermidis ope	226	6	3.8	269	16	AAAR82017	Variant lipase of
154	6	3.8	163	18	AAAM13651	Human CTIA-8. Hom	227	6	3.8	269	16	AAAR82018	Variant lipase of
155	6	3.8	163	21	AAAB03835	Human cytotoxic T	228	6	3.8	269	16	AAAR82019	Variant lipase of
156	6	3.8	163	21	AAAY81986	Human CTIA-8 prote	229	6	3.8	269	16	AAAR82020	Variant lipase of
157	6	3.8	163	22	AAAU04954	Human Interleukin	230	6	3.8	269	16	AAAR82021	Variant lipase of
											18	AAW211155	Lipolytic enzyme,

231	269	18	AAW21156	Lipolytic enzyme,	304	6	3.8	275	18	AAW21193	Lipolytic enzyme,E
232	269	18	AAW21144	Lipolytic enzyme,	305	6	3.8	275	18	AAW21196	Lipolytic enzyme,E
233	269	18	AAW21145	Lipolytic enzyme,	306	6	3.8	275	18	AAW21197	Lipolytic enzyme,E
234	269	18	AAW21146	Lipolytic enzyme,	307	6	3.8	275	18	AAW21180	Lipolytic enzyme,E
235	269	18	AAW21147	Lipolytic enzyme,	308	6	3.8	275	18	AAW21181	Lipolytic enzyme,E
236	269	18	AAW21148	Lipolytic enzyme,	309	6	3.8	275	18	AAW21182	Lipolytic enzyme,E
237	269	18	AAW21149	Lipolytic enzyme,	310	6	3.8	275	18	AAW21183	Lipolytic enzyme,E
238	269	18	AAW21150	Lytic enzyme, N94K	311	6	3.8	275	18	AAW21184	Lipolytic enzyme,E
239	269	18	AAW21151	Lipolytic enzyme,	312	6	3.8	275	18	AAW21176	Lipolytic enzyme,E
240	269	18	AAW21152	Lipolytic enzyme,	313	6	3.8	275	18	AAW21177	Lipolytic enzyme,E
241	269	18	AAW21153	Lipolytic enzyme,	314	6	3.8	275	18	AAW21178	Lipolytic enzyme,E
242	269	18	AAW21154	Lipolytic enzyme,	315	6	3.8	275	18	AAW21179	Lipolytic enzyme,E
243	269	18	AAW21153	Lipolytic enzyme,	316	6	3.8	275	18	AAW21173	Lipolytic enzyme,E
244	269	18	AAW21134	Lipolytic enzyme,	317	6	3.8	275	18	AAW21167	Lipolytic enzyme,E
245	269	18	AAW21135	Lipolytic enzyme,	318	6	3.8	275	18	AAW21168	Lipolytic enzyme,E
246	269	18	AAW21136	Lipolytic enzyme,	319	6	3.8	275	18	AAW21169	Lipolytic enzyme,E
247	269	18	AAW21137	Lipolytic enzyme,	320	6	3.8	275	18	AAW21162	Lipolytic enzyme,
248	269	18	AAW21138	Lipolytic enzyme,	321	6	3.8	275	18	AAW21158	Lipolytic enzyme,
249	269	18	AAW21139	Lipolytic enzyme,	322	6	3.8	275	22	AAW21158	Mouse ageing inhib
250	269	18	AAW21140	Lytic enzyme N94K+	323	6	3.8	275	18	AAW21199	Lipolytic enzyme,
251	269	18	AAW21141	Lipolytic enzyme,	324	6	3.8	276	18	AAW21194	Lipolytic enzyme,E
252	269	18	AAW21142	Lipolytic enzyme,	325	6	3.8	276	18	AAW21195	Lipolytic enzyme,E
253	269	18	AAW21143	Lipolytic enzyme,	326	6	3.8	276	18	AAW21195	Lipolytic enzyme,E
254	269	18	AAW21121	Lipolytic enzyme,	327	6	3.8	276	18	AAW21171	Lipolytic enzyme,E
255	269	18	AAW21122	Lipolytic enzyme,	328	6	3.8	276	18	AAW21172	Lipolytic enzyme,E
256	269	18	AAW21123	Lipolytic enzyme,	329	6	3.8	276	18	AAW21174	Lipolytic enzyme,E
257	269	18	AAW21124	Lipolytic enzyme,	330	6	3.8	276	18	AAW21170	Lipolytic enzyme,E
258	269	18	AAW21125	Lipolytic enzyme,	331	6	3.8	276	18	AAW21161	Lipolytic enzyme,
259	269	18	AAW21126	Lipolytic enzyme,	332	6	3.8	276	18	AAW21163	Lipolytic enzyme,E
260	269	18	AAW21127	Lipolytic enzyme,	333	6	3.8	276	18	AAW21164	Lipolytic enzyme,E
261	269	18	AAW21128	Lipolytic enzyme,	334	6	3.8	276	18	AAW21165	Lipolytic enzyme,E
262	269	18	AAW21129	Lipolytic enzyme,	335	6	3.8	276	18	AAW21157	Lipolytic enzyme,
263	269	18	AAW21130	Lipolytic enzyme,	336	6	3.8	276	18	AAW21159	Lipolytic enzyme,
264	269	18	AAW21131	Lipolytic enzyme,	337	6	3.8	277	18	AAW21200	Lipolytic enzyme,
265	269	18	AAW21132	Lipolytic enzyme,	338	6	3.8	277	18	AAW21198	Lipolytic enzyme,E
266	269	18	AAW21110	Lipolytic enzyme,	339	6	3.8	277	18	AAW21160	Lipolytic enzyme,
267	269	18	AAW21111	Lipolytic enzyme,	340	6	3.8	278	18	AAW21166	Lipolytic enzyme,E
268	269	18	AAW21112	Lipolytic enzyme,	341	6	3.8	279	20	AAW21157	Staphylococcus aur
269	269	18	AAW21113	Lipolytic enzyme,	342	6	3.8	279	20	AAW21157	S. aureus spo0J2 p
270	269	18	AAW21114	Lipolytic enzyme,	343	6	3.8	281	18	AAW21189	Lipolytic enzyme,E
271	269	18	AAW21115	Lipolytic enzyme,	344	6	3.8	281	18	AAW21190	Lytic enzyme, ElApp
272	269	18	AAW21116	Lipolytic enzyme,	345	6	3.8	281	18	AAW21191	Lipolytic enzyme,E
273	269	18	AAW21117	Lipolytic enzyme,	346	6	3.8	281	21	AAW21191	Arabidopsis thalia
274	269	18	AAW21118	Lipolytic enzyme,	347	6	3.8	281	21	AAW21191	Arabidopsis thalia
275	269	18	AAW21119	Lipolytic enzyme,	348	6	3.8	283	15	AAW21191	Theileria buffeli
276	269	18	AAW21120	Lipolytic enzyme,	349	6	3.8	285	13	AAW21191	Humicola lanuginos
277	269	18	AAW21109	Lipolytic enzyme,	350	6	3.8	286	21	AAW21191	Gene 2 human secre
278	269	18	AAW21110	Lipolytic enzyme,	351	6	3.8	286	21	AAW21191	Arabidopsis thalia
279	269	18	AAW21101	Lipolytic enzyme,	352	6	3.8	287	21	AAW21196	Humicola lanuginos
280	269	18	AAW21102	Lipolytic enzyme,	353	6	3.8	288	13	AAW21196	C-7 hydroxycephem
281	269	18	AAW21103	Lipolytic enzyme,	354	6	3.8	288	13	AAW21196	Humicola lanuginos
282	269	18	AAW21104	Lipolytic enzyme,	355	6	3.8	289	13	AAW21196	Humicola lanuginos
283	269	18	AAW21105	Lipolytic enzyme,	356	6	3.8	289	13	AAW21196	Product of Humicol
284	269	18	AAW21106	Lipolytic enzyme,	357	6	3.8	289	18	AAW21196	Humicola lanuginos
285	269	18	AAW21107	Lipolytic enzyme,	358	6	3.8	290	13	AAW21196	Humicola lanuginos
286	269	18	AAW21108	Lipolytic enzyme,	359	6	3.8	291	13	AAW21196	Humicola lanuginos
287	269	18	AAW21109	Lipolytic enzyme,	360	6	3.8	291	13	AAW21196	Humicola lanuginos
288	269	18	AAW21109	Wild type H. lanug	361	6	3.8	291	13	AAW21196	Humicola lanuginos
289	269	18	AAW21093	Lipolytic enzyme,	362	6	3.8	291	13	AAW21196	Humicola lanuginos
290	269	18	AAW21094	Lipolytic enzyme,	363	6	3.8	291	13	AAW21196	Humicola lanuginos
291	269	18	AAW21095	Lipolytic enzyme,	364	6	3.8	291	13	AAW21196	Humicola lanuginos
292	269	18	AAW21096	Lipolytic enzyme,	365	6	3.8	291	13	AAW21196	Humicola lanuginos
293	269	18	AAW21097	Lipolytic enzyme,	366	6	3.8	291	13	AAW21196	Humicola lanuginos
294	269	18	AAW21098	Lipolytic enzyme,	367	6	3.8	291	13	AAW21196	Humicola lanuginos
295	269	18	AAW21098	Lipase of Humicola	368	6	3.8	291	13	AAW21196	Humicola lanuginos
296	269	18	AAW21098	S. epidermidis ope	369	6	3.8	291	13	AAW21196	Humicola lanuginos
297	269	18	AAW21098	Humanised ID10 ant	370	6	3.8	291	13	AAW21196	Humicola lanuginos
298	269	18	AAW21098	Lipolytic enzyme,	371	6	3.8	291	13	AAW21196	Humicola lanuginos
299	269	18	AAW21098	Lipolytic enzyme,	372	6	3.8	291	13	AAW21196	Humicola lanuginos
300	269	18	AAW21098	Human protein HPI0	373	6	3.8	291	13	AAW21196	H. lanuginosa lipa
301	269	18	AAW21185	Lipolytic enzyme,E	374	6	3.8	291	13	AAW21196	H. lanuginosa lipa
302	269	18	AAW21186	Lipolytic enzyme,E	375	6	3.8	291	13	AAW21196	H. lanuginosa lipa
303	269	18	AAW21187	Lipolytic enzyme,E	376	6	3.8	291	13	AAW21196	Sequence of pre-pr
				Lipolytic enzyme,E				291	15	AAW21196	Variant lipase D77

377	6	3.8	291	15	AAW65445	Variant lipase D96	450	6	3.8	294	18	AAW23216	Product of Humicol
378	6	3.8	291	15	AAW65446	Variant lipase D96	451	6	3.8	294	18	AAW19646	Humicola lanuginos
379	6	3.8	291	15	AAW65447	Variant lipase D96	452	6	3.8	294	18	AAW63449	Human breast cancer
380	6	3.8	291	15	AAW65448	Variant lipase D96	453	6	3.8	295	22	AAW73666	Mouse ageing inhib
381	6	3.8	291	15	AAW65449	Variant lipase E56	454	6	3.8	305	21	AAW33465	Arabidopsis thalia
382	6	3.8	291	15	AAW65428	Variant lipase L25	455	6	3.8	306	22	AAW93028	C glutamicum prote
383	6	3.8	291	15	AAW65429	Variant lipase L20	456	6	3.8	306	22	AAW71704	Human olfactory re
384	6	3.8	291	15	AAW65430	Variant lipase L20	457	6	3.8	306	22	AAW72470	Human OR-like poly
385	6	3.8	291	15	AAW65431	Variant lipase L20	458	6	3.8	306	22	AAW79226	Corynebacterium gl
386	6	3.8	291	15	AAW65432	Variant lipase L20	459	6	3.8	312	22	AAW72138	Human olfactory re
387	6	3.8	291	15	AAW65433	Variant lipase L20	460	6	3.8	316	22	AAW42210	Human polypeptide
388	6	3.8	291	15	AAW65434	Variant lipase L20	461	6	3.8	317	20	AAW29262	Amino acid sequenc
389	6	3.8	291	15	AAW65435	Variant lipase L20	462	6	3.8	322	20	AAW31601	Escherichia coli b
390	6	3.8	291	15	AAW65436	Variant lipase L20	463	6	3.8	325	21	AAW09462	Arabidopsis thalia
391	6	3.8	291	15	AAW65437	Variant lipase L20	464	6	3.8	325	21	AAW44074	Arabidopsis thalia
392	6	3.8	291	15	AAW65438	Variant lipase L20	465	6	3.8	325	21	AAW48292	Arabidopsis thalia
393	6	3.8	291	15	AAW65439	Variant lipase L20	466	6	3.8	326	21	AAW33464	Arabidopsis thalia
394	6	3.8	291	15	AAW65440	Variant lipase L20	467	6	3.8	331	21	AAW17625	Arabidopsis thalia
395	6	3.8	291	15	AAW65414	Variant lipase L20	468	6	3.8	331	21	AAW45478	Arabidopsis thalia
396	6	3.8	291	15	AAW65441	Variant lipase L20	469	6	3.8	337	22	AAW05131	Human odorant rece
397	6	3.8	291	15	AAW65442	Variant lipase D96	470	6	3.8	337	22	AAW72130	Human olfactory re
398	6	3.8	291	15	AAW65412	Variant lipase D96	471	6	3.8	340	19	AAW81594	Protein encoded by
399	6	3.8	291	15	AAW65413	Variant lipase D96	472	6	3.8	340	22	AAW04820	Micromonospora eve
400	6	3.8	291	15	AAW65414	Variant lipase D96	473	6	3.8	341	21	AAW97257	Human IgG1 FC-CTIA
401	6	3.8	291	15	AAW65415	Variant lipase D96	474	6	3.8	341	21	AAW97260	Human IgG1 FC-CTIA
402	6	3.8	291	15	AAW65416	Variant lipase D96	475	6	3.8	343	22	AAW67575	Amino acid sequenc
403	6	3.8	291	15	AAW65417	Variant lipase D96	476	6	3.8	345	21	AAW97256	Human IgG1 FC-HSV1
404	6	3.8	291	15	AAW65418	Variant lipase L25	477	6	3.8	345	21	AAW97259	Human IgG1 FC-HSV1
405	6	3.8	291	15	AAW65419	Variant lipase L25	478	6	3.8	348	17	AAW01437	Murine cyclin-depe
406	6	3.8	291	15	AAW65420	Variant lipase L25	479	6	3.8	351	21	AAW63092	Human secreted pro
407	6	3.8	291	15	AAW65421	Variant lipase L25	480	6	3.8	352	22	AAW43588	Human polypeptide
408	6	3.8	291	15	AAW65422	Variant lipase L25	481	6	3.8	354	21	AAW42435	Human ORFX ORF2199
409	6	3.8	291	15	AAW65423	Variant lipase L25	482	6	3.8	354	22	AAW39840	Human polypeptide
410	6	3.8	291	15	AAW65424	Variant lipase L25	483	6	3.8	355	22	AAW41626	Human polypeptide
411	6	3.8	291	15	AAW65425	Variant lipase L25	484	6	3.8	356	21	AAW04120	Arabidopsis thalia
412	6	3.8	291	15	AAW65426	Variant lipase L25	485	6	3.8	359	22	AAW67841	Amino acid sequenc
413	6	3.8	291	15	AAW65427	Variant lipase L25	486	6	3.8	360	22	AAW41900	Human polypeptide
414	6	3.8	291	15	AAW65396	Variant lipase E56	487	6	3.8	361	21	AAW08538	A human G-protein
415	6	3.8	291	15	AAW65397	Variant lipase E56	488	6	3.8	361	21	AAW44662	Human 14273 G-prot
416	6	3.8	291	15	AAW65398	Variant lipase E56	489	6	3.8	361	21	AAW44815	Human 14273 G-prot
417	6	3.8	291	15	AAW65399	Variant lipase E56	490	6	3.8	361	21	AAW44817	Murine 14273 G-pro
418	6	3.8	291	15	AAW65400	Variant lipase E56	491	6	3.8	361	22	AAW04564	Human G-protein co
419	6	3.8	291	15	AAW65401	Variant lipase E56	492	6	3.8	362	21	AAW44814	Human 14273 G-prot
420	6	3.8	291	15	AAW65402	Variant lipase E56	493	6	3.8	365	4	AAW30603	Sequence encoded b
421	6	3.8	291	15	AAW65403	Variant lipase E56	494	6	3.8	365	10	AAW94144	prochymosin. AAP
422	6	3.8	291	15	AAW65404	Variant lipase E56	495	6	3.8	365	11	AAW05080	Sequence of calf p
423	6	3.8	291	15	AAW65405	Variant lipase E56	496	6	3.8	370	21	AAW04119	Arabidopsis thalia
424	6	3.8	291	15	AAW65406	Variant lipase E56	497	6	3.8	375	5	AAW40078	Sequence encoded b
425	6	3.8	291	15	AAW65407	Variant lipase D96	498	6	3.8	376	17	AAW90341	pJG4-5-CDK-BP clon
426	6	3.8	291	15	AAW65408	Variant lipase D96	499	6	3.8	377	22	AAW91069	C glutamicum prote
427	6	3.8	291	15	AAW65395	Wild type lipase u	500	6	3.8	378	21	AAW53266	Phagemid pFab-SP40
428	6	3.8	291	15	AAW65409	Variant lipase D96	501	6	3.8	379	4	AAW30013	Sequence encoded b
429	6	3.8	291	15	AAW65410	Variant lipase D96	502	6	3.8	380	3	AAW20038	pre-prorennin-A pr
430	6	3.8	291	15	AAW65411	Variant lipase D96	503	6	3.8	381	4	AAW30086	Sequence encoded b
431	6	3.8	291	16	AAW81990	Lipase of Humicola	504	6	3.8	381	4	AAW30446	Sequence encoded b
432	6	3.8	291	17	AAW24074	Lipolase. Bacillu	505	6	3.8	381	5	AAW40218	Sequence of rennin
433	6	3.8	291	17	AAW88472	1,3-Lipase Lipolas	506	6	3.8	381	5	AAW40559	Sequence of a poly
434	6	3.8	291	18	AAW32790	Lipase enzyme used	507	6	3.8	381	13	AAW20730	prochymosin (prote
435	6	3.8	291	18	AAW32791	SEQ ID NO:2 from s	508	6	3.8	384	19	AAW60212	Mutan binding doma
436	6	3.8	291	18	AAW32623	Lipase for use in	509	6	3.8	386	19	AAW72064	HSV-2 strain SB5 C
437	6	3.8	291	18	AAW32622	Lipase for use in	510	6	3.8	386	21	AAW74678	Neisseria meningit
438	6	3.8	291	18	AAW13560	Humicola lanuginos	511	6	3.8	386	21	AAW74679	Neisseria meningit
439	6	3.8	291	18	AAW13561	Humicola lanuginos	512	6	3.8	387	10	AAW93672	Sequence from acyB
440	6	3.8	291	18	AAW13556	Humicola lanuginos	513	6	3.8	390	21	AAW74677	Neisseria gonorrhoe
441	6	3.8	291	18	AAW13557	Humicola lanuginos	514	6	3.8	390	22	AAU00536	Bovine chymosin po
442	6	3.8	291	18	AAW11845	Product of Humicol	515	6	3.8	391	21	AAW53267	Phagemid ph8 wild
443	6	3.8	291	18	AAW14423	Humicola lanuginos	516	6	3.8	397	19	AAW72086	HSV-2 strain SB5 C
444	6	3.8	291	19	AAW75708	Humicola lanuginos	517	6	3.8	399	19	AAW81590	Protein encoded by
445	6	3.8	291	20	AAW03661	1,3-specific lipas	518	6	3.8	404	19	AAW68408	Aufjeszky's disease
446	6	3.8	292	21	AAW94666	Prepro Humicola la	519	6	3.8	408	19	AAW72138	HSV-2 strain SB5 C
447	6	3.8	292	18	AAW13562	Humicola lanuginos	520	6	3.8	408	19	AAW71465	Glycosyltransferas
448	6	3.8	292	18	AAW13558	Humicola lanuginos	521	6	3.8	408	21	AAW67549	A. orientalis glyc
449	6	3.8	293	22	AAU00036	Wheat Starch synth	522	6	3.8	408	22	AAW40414	Human polypeptide

523	6	3.8	408	22	AAB62450	S. cerevisiae YOL0	596	3.8	583	21	AAB19072	Amino acid sequenc
524	6	3.8	410	20	AAW93577	Human APORP protei	597	3.8	583	22	AAW90782	Human shear stress
525	6	3.8	410	21	AAB13565	Streptomyces globi	598	3.8	584	20	AAW34931	Amino acid sequenc
526	6	3.8	412	20	AAW94998	Serine/threonine p	599	3.8	584	22	AAB66393	Chlamydia pneumoni
527	6	3.8	415	21	AAW77196	S. venezuelae macr	600	3.8	596	20	AAW07040	Breast cancer asso
528	6	3.8	416	21	AAB18654	Amino acid sequenc	601	3.8	609	21	AAW31959	Arabidopsis thalia
529	6	3.8	416	21	AAW67218	Macrolide ring mod	602	3.8	616	22	AAB88551	Haemophilus influe
530	6	3.8	417	22	AAU01257	E. coli alanine-va	603	3.8	624	21	AAG31958	Arabidopsis thalia
531	6	3.8	418	21	AAG33463	Arabidopsis thalia	604	3.8	640	22	AAW31958	Human protein sequ
532	6	3.8	419	22	AAW75601	Gene 45 human secr	605	3.8	659	21	AAG31957	Arabidopsis thalia
533	6	3.8	419	22	AAW75602	Human secreted pro	606	3.8	668	22	AAW94411	Human protein sequ
534	6	3.8	424	20	AAW35407	Chlamydia pneumoni	607	3.8	708	20	AAW13452	Amino acid sequenc
535	6	3.8	424	22	AAW38975	Human polypeptide	608	3.8	721	19	AAW72211	HSV-2 strain S95 C
536	6	3.8	424	22	AAW89887	C. glutamicum SPT	609	3.8	738	22	AAW00978	Human bone marrow
537	6	3.8	424	22	AAW87854	C. glutamicum prote	610	3.8	750	19	AAW72176	HSV-2 strain S95 C
538	6	3.8	424	22	AAW76809	Corynebacterium gl	611	3.8	773	10	AAW91553	Membrane-bound ald
539	6	3.8	433	20	AAW42277	Salmonella dublin	612	3.8	781	21	AAW52022	A. fulgidus AF0497
540	6	3.8	438	20	AAW37229	Chlamydia trachoma	613	3.8	781	21	AAW51651	A. fulgidus AF0497
541	6	3.8	439	22	AAW73634	Human colon cancer	614	3.8	790	18	AAW31528	Human anti-apoptot
542	6	3.8	441	21	AAW42997	Human ORFX ORF2761	615	3.8	790	21	AAW41923	Arabidopsis thalia
543	6	3.8	441	22	AAW92914	Human protein sequ	616	3.8	790	22	AAW47406	Human A20. Homo s
544	6	3.8	441	22	AAW94700	Human protein sequ	617	3.8	791	22	AAW00952	Human bone marrow
545	6	3.8	441	22	AAW64382	Amino acid sequenc	618	3.8	794	21	AAW41922	Arabidopsis thalia
546	6	3.8	443	21	AAW75478	Neisseria meningit	619	3.8	802	22	AAW95231	Human protein sequ
547	6	3.8	447	22	AAW70714	S. cerevisiae apopt	620	3.8	803	15	AAW60882	P. abyssi DNA poly
548	6	3.8	450	10	AAW94370	Sequence encoded b	621	3.8	806	15	AAW55692	hRNP U protein.
549	6	3.8	450	21	AAW29195	Arabidopsis thalia	622	3.8	814	21	AAW41921	Arabidopsis thalia
550	6	3.8	451	21	AAW29194	Arabidopsis thalia	623	3.8	815	22	AAW25875	Human protein sequ
551	6	3.8	452	12	AAW13119	Phenylalanine hydr	624	3.8	817	22	AAW38657	Human polypeptide
552	6	3.8	452	18	AAW25788	Human phenylalanin	625	3.8	819	22	AAW40442	Human polypeptide
553	6	3.8	452	20	AAW55893	Human phenylalanin	626	3.8	819	22	AAW40443	Human polypeptide
554	6	3.8	452	21	AAW78593	Human phenylalanin	627	3.8	857	22	AAW23698	Human EST encoded
555	6	3.8	455	22	AAW74018	Human colon cancer	628	3.8	863	22	AAW38656	Human polypeptide
556	6	3.8	456	21	AAW20758	Arabidopsis thalia	629	3.8	877	16	AAW82657	Human mclur3. Hom
557	6	3.8	456	21	AAW50449	Arabidopsis thalia	630	3.8	879	16	AAW64252	Human mclur3. Hom
558	6	3.8	458	10	AAW94376	BamHI/SalI insert	631	3.8	880	21	AAW13632	C. trachomatis pmp
559	6	3.8	463	18	AAW19541	Soybean thiol prot	632	3.8	880	22	AAW83200	Protein encoded by
560	6	3.8	464	18	AAW19542	Soybean thiol prot	633	3.8	887	16	AAW65492	Pyruvate dehydroge
561	6	3.8	467	22	AAW50896	Human DR4. Homo s	634	3.8	998	20	AAW76809	N. gonorrhoea pilc
562	6	3.8	468	19	AAW64483	Human DR4 protein.	635	3.8	1000	21	AAW58914	Fumonisin esterase
563	6	3.8	468	20	AAW31802	Human death recept	636	3.8	1000	21	AAW68850	A fusion of aminop
564	6	3.8	468	20	AAW93609	Human DR4 protein.	637	3.8	1003	21	AAW24379	E. tenella cgm de
565	6	3.8	468	21	AAW72022	Human Death Domain	638	3.8	1014	22	AAW73665	Mouse ageing inhib
566	6	3.8	468	21	AAW72023	Human Death Domain	639	3.8	1026	17	AAW97374	Phage T4 tail fibr
567	6	3.8	468	21	AAW08546	Amino acid sequenc	640	3.8	1026	20	AAW94678	Bacteriophage T4 t
568	6	3.8	468	21	AAW01339	TNF-related apopt	641	3.8	1026	20	AAW92358	Bacteriophage T4 t
569	6	3.8	468	22	AAW49241	Human DR4 protein.	642	3.8	1026	22	AAW35154	Bacteriophage T4 g
570	6	3.8	481	21	AAW20757	Arabidopsis thalia	643	3.8	1029	13	AAW26788	Sequence encoded b
571	6	3.8	481	21	AAW50448	Arabidopsis thalia	644	3.8	1029	20	AAW76807	N. gonorrhoea pilc
572	6	3.8	482	21	AAW96786	Soybean sucrose no	645	3.8	1043	22	AAW73664	Mouse ageing inhib
573	6	3.8	486	22	AAW93892	Human protein sequ	646	3.8	1098	22	AAW25720	Human protein sequ
574	6	3.8	490	16	AAW72566	Penicillium purpur	647	3.8	1138	11	AAW06461	Human protein sequ
575	6	3.8	496	15	AAW57349	Human cartilage ma	648	3.8	1138	14	AAW37213	BtPGS1245 protoxin
576	6	3.8	497	20	AAW99362	C. trachomatis Cys	649	3.8	1138	14	AAW37214	B.t. toxin HD511.
577	6	3.8	500	20	AAW37360	Protein involved i	650	3.8	1138	15	AAW46225	Bacillus thuringie
578	6	3.8	503	21	AAW81750	Streptococcus pneu	651	3.8	1138	15	AAW46226	Bacillus thuringie
579	6	3.8	509	21	AAW04118	Arabidopsis thalia	652	3.8	1205	21	AAW58916	Fumonisin esterase
580	6	3.8	511	21	AAW32525	S. lavendulae Mmcl	653	3.8	1205	21	AAW68852	Fumonisin esterase
581	6	3.8	512	21	AAW74680	Neisseria gonorrhe	654	3.8	1256	22	AAW84855	A fusion of aminop
582	6	3.8	512	21	AAW74682	Neisseria meningit	655	3.8	1310	22	AAW40851	Murine nephrin pro
583	6	3.8	512	21	AAW74682	Neisseria meningit	656	3.8	1350	22	AAW02944	Human polypeptide
584	6	3.8	525	20	AAW94289	BAA-ESPI fusion pr	657	3.8	1445	20	AAW93595	Angiotensin conver
585	6	3.8	527	17	AAW91277	Exophiala spinifer	658	3.8	1464	14	AAW44192	O. longistaminata
586	6	3.8	527	20	AAW94287	E. spinifera fumon	659	3.8	1464	14	AAW42054	Rat NMDA receptor
587	6	3.8	535	21	AAW08851	A murine heparanas	660	3.8	1464	15	AAW66039	Glutamic acid rece
588	6	3.8	545	20	AAW33830	Oleolin-spacer-Met	661	3.8	1464	15	AAW55529	Human N-methyl-D-a
589	6	3.8	548	21	AAW20756	Arabidopsis thalia	662	3.8	1464	16	AAW80970	Human NMDA R2A rec
590	6	3.8	548	21	AAW50447	Arabidopsis thalia	663	3.8	1464	20	AAW85976	Human excitatory a
591	6	3.8	549	22	AAW91013	C. glutamicum prote	664	3.8	1464	21	AAW26215	Human N-methyl-D-a
592	6	3.8	556	20	AAW13453	Amino acid sequenc	665	3.8	1464	21	AAW56113	Human NMDAR2A subu
593	6	3.8	559	20	AAW34475	Porphorymonas ging	666	3.8	1468	21	AAW45856	Arabidopsis thalia
594	6	3.8	571	19	AAW74889	Human secreted pro	667	3.8	1481	21	AAW10482	S. aggregatum PKS
595	6	3.8	582	20	AAW34351	Porphorymonas ging	668	3.8	1523	21	AAW45855	Arabidopsis thalia

669 6 3.8 1600 22 AAG84976 Shrimp white spot  
 670 6 3.8 1636 21 AAG45854 Arabidopsis thalia  
 671 6 3.8 1670 21 AAG31673 Arabidopsis thalia  
 672 6 3.8 1753 21 AAG31672 Arabidopsis thalia  
 673 6 3.8 1754 21 AAG52410 Arabidopsis thalia  
 674 6 3.8 1769 21 AAG52409 Arabidopsis thalia  
 675 6 3.8 1804 21 AAG52408 Arabidopsis thalia  
 676 6 3.8 1816 21 AAG31671 Arabidopsis thalia  
 677 6 3.8 2115 19 AAW59276 Rubella virus RA27  
 678 6 3.8 2194 22 AAM40114 Human polypeptide  
 679 6 3.8 2205 16 AAR79048 Infectious rubella  
 680 6 3.8 2442 21 AAR77575 Human cytoskeletal  
 681 6 3.8 3164 16 AAR94346 Hepatitis GB virus  
 682 6 3.8 3567 14 AAR44431 eryA region polype  
 683 6 3.8 3587 14 AAR34713 Bacillus subtilis  
 684 6 3.8 3588 14 AAR34712 Bacillus subtilis

## ALIGNMENTS

## RESULT 1

AA53891  
 ID AAY53891 standard; Protein; 160 AA.

AC AAY53891;

DT 13-MAR-2000 (first entry)

DE Partial amino acid sequence of human Interleukin-22.

XX Human; Interleukin-22; IL-22; IL-21; Immune system disorder;  
 KW Immune cell chemotaxis; haematopoietic cell disorder;  
 KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;  
 KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;  
 KW inflammation; hyperproliferative disorder; tissue regeneration;  
 KW embryonic stem cell differentiation; embryonic stem cell proliferation;  
 KW haematopoietic lineage; allergic asthma.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 26 /note= "potential N-linked glycosylation site"

FT Domain 57..64 /note= "conserved domain I"

FT Domain 72..77 /note= "conserved domain II"

FT Domain 99..105 /note= "conserved domain III"

FT Domain 121..128 /note= "conserved domain IV"

FT Modified-site 139 /note= "potential N-linked glycosylation site"

FT

XX WO9961617-A1.

XX 02-DEC-1999.

XX 27-MAY-1999; 99WO-US11644.

XX 29-MAY-1998; 98US-0087340.

PR 10-SEP-1998; 98US-0099805.

PR 30-APR-1999; 99US-0131965.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Ebner R;

XX WPI; 2000-072622/06.

DR N-PSDB; AAZ36835.

XX Novel polynucleotides used to develop products for treating e.g. immune

PT disorders, blood disorders, autoimmune disorders, allergies,  
 PT inflammation, hyperproliferative disorders or infections -  
 XX Claim 27; Fig 2A-B; 170pp; English.

XX The present sequence represents a partial human interleukin-22 (IL-22)  
 CC protein. The specification also describes IL-21 polynucleotides and  
 CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library  
 CC of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating  
 CC deficiencies or disorders of the immune system, by activating or  
 CC inhibiting the proliferation, differentiation, or mobilization  
 CC (chemotaxis) of immune cells, treating or detecting deficiencies or  
 CC disorders of haematopoietic cells, to modulate haemostatic or  
 CC thrombolytic activity, in treating or detecting autoimmune disorders,  
 CC treating asthma (particularly allergic asthma) or other respiratory  
 CC problems, to treat and/or prevent organ rejection or graft-versus-host  
 CC disease (GVHD), to modulate inflammation, to treat or detect  
 CC hyperproliferative disorders, to treat or detect infectious agents, to  
 CC differentiate, proliferate and attract cells, leading to the  
 CC regeneration of tissues, IL-21 and IL-22 may also increase or decrease  
 CC the differentiation or proliferation of embryonic stem cells and  
 CC haematopoietic lineage, may be used to modulate mammalian  
 CC characteristics.

XX Sequence 160 AA;

Query Match 100.0%; Score 160; DB 21; Length 160;

Best Local Similarity 100.0%; Pred. No. 7.2e-156;

Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSARARAVLSAFHHTLQLGPREQARNASCAGGPADRRFPPTNLRSPWPAYRISYDP 60  
 Db 1 nsararavlsafhhtlqlgpreqarnascaggpadrfrfptnlrsvspwayrisydp 60

Qy 61 ARYPRLPEAYCLRGCLTGLFGEDVFRSAPVYMTVLLRPPACAGGRSVTEAYVT 120  
 Db 61 aryprylpeayclrgcltglfgeedvfrsapyvmtvllrrtpacaggrsvteayvt 120

Qy 121 IPVGCTCVPEPKDADSDINSIDKQAKLLGPNADAPG 160  
 Db 121 ipvgctcvpepkdadsinsidkqakllgpnadapg 160

RESULT 2

AA53893

ID AAY53893 standard; Protein; 173 AA.

XX AAY53893;

DT 13-MAR-2000 (first entry)

DE Partial amino acid sequence of human Interleukin-22.

XX Human; Interleukin-22; IL-22; IL-21; Immune system disorder;  
 KW Immune cell chemotaxis; haematopoietic cell disorder;  
 KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;  
 KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;  
 KW inflammation; hyperproliferative disorder; tissue regeneration;  
 KW embryonic stem cell differentiation; embryonic stem cell proliferation;  
 KW haematopoietic lineage; allergic asthma.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Domain 18..23 /note= "conserved domain VI"

FT Modified-site 39..41 /note= "Asn39 is a potential N-linked glycosylation site"

FT Domain 60..65 /note= "conserved domain VII"

FT Domain 69..77 /note= "conserved domain I"

FT

FT Domain 85..90  
 FT /note= "conserved domain II"  
 FT 112..118  
 FT Domain  
 FT /note= "conserved domain III"  
 FT 134..141  
 FT Domain  
 FT /note= "conserved domain IV"  
 FT 152..154  
 FT Modified-site  
 FT /note= "Asn152 is a potential N-linked glycosylation site"  
 FT XX  
 FT 27-MAY-1999; 99WO-US11644.  
 FT XX  
 FT 29-MAY-1998; 98US-0087340.  
 PR 10-SEP-1998; 98US-0099805.  
 PR 30-APR-1999; 99US-0131965.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Ruben SM, Ebner R;  
 XX  
 DR WPI: 2000-072622/06.  
 DR N-PSDB; AA236837.  
 XX  
 PT Novel polynucleotides used to develop products for treating e.g. immune disorders, blood disorders, autoimmune disorders, allergies, inflammation, hyperproliferative disorders or infections -  
 PT  
 XX Disclosure; Fig 8; 170pp; English.  
 XX  
 CC The present sequence represents a partial human interleukin-22 (IL-22) protein. The specification also describes IL-21 polynucleotides and polypeptides. The IL-21 polynucleotide was isolated from a cDNA library of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization of haematopoietic cells, treating or detecting deficiencies (chemotaxis) of immune cells, treating or detecting deficiencies of haematopoietic cells, to modulate haemostatic or thrombolytic activity, in treating or detecting autoimmune disorders, treating asthma (particularly allergic asthma) or other respiratory problems, to treat and/or prevent organ rejection or graft-versus-host disease (GVHD), to modulate inflammation, to treat or detect hyperproliferative disorders, to treat or detect infectious agents, to differentiate, proliferate and attract cells, leading to the regeneration of tissues, CC IL-21 and IL-22 may also increase or decrease the differentiation or proliferation of embryonic stem cells and haematopoietic lineage, may be used to modulate mammalian characteristics.  
 CC  
 XX Sequence 173 AA;

Query Match 95.6%; Score 153; DB 21; Length 173;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-148;  
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 VLSAFHHTLQLGPREQARNASCAGGPPADRRFRPPNLRSPWYRISYDPRYRL 67  
 Db 21 VLSAFHHTLQLGPREQARNASCAGGPPADRRFRPPNLRSPWYRISYDPRYRL 80  
 QY 68 PEAYCLRCGLTGLFGEDVRFSPVYMTVLRTPACAGGRSVYTEAYVTIPVGCCTC 127  
 Db 81 PEAYCLRCGLTGLFGEDVRFSPVYMTVLRTPACAGGRSVYTEAYVTIPVGCCTC 140  
 QY 128 VPEPEKDADSIINSIDKQAKLLIGPNDAPAGP 160  
 Db 141 VPEPEKDADSIINSIDKQAKLLIGPNDAPAGP 173

RESULT 3  
 AAY70656

ID AAY70656 standard; Protein; 185 AA.  
 XX  
 AC AAY70656;  
 XX  
 DT 18-JUL-2000 (first entry)  
 XX  
 DE Mature human transforming growth factor beta-9, ztgbeta-9 protein-3.  
 XX  
 KW Human transforming growth factor beta-9; ztgbeta-9;  
 KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
 KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
 KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
 KW antiviral; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200015798-A2.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PF 17-SEP-1999; 99WO-US21677.  
 XX  
 PR 17-SEP-1998; 98US-0154817.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Presnell SR, Taft DW, Foley KP;  
 XX  
 DR WPI: 2000-271436/23.  
 DR N-PSDB; AA252195.  
 XX  
 PT Polynucleotides encoding a novel transforming growth factor beta-9 polypeptide, designated ztgbeta-9, useful as an antiviral and antiproliferative agent -  
 PT  
 XX  
 PS Claim 6; Page 85-86; 97pp; English.  
 XX  
 CC The present sequence is the mature human transforming growth factor beta-9, designated ztgbeta-9. This is a mature sequence excluding the signal sequence extending from amino acid 18 to and including amino acid 202 of ztgbeta-9. Human ztgbeta-9 was isolated from an arrayed CC pituitary gland cDNA plasmid library by PCR screening. This can be used CC to treat a variety of neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease, CC Parkinson's disease and peripheral neuropathies, or demyelinating diseases including multiple sclerosis. ztgbeta-9 peptides have CC antiviral activity and may also be used to regulate the proliferation, CC differentiation and apoptosis of neurons, glial cells, lymphocytes, CC hematopoietic cells and stromal cells.  
 XX  
 SQ Sequence 185 AA;  
 Query Match 95.6%; Score 153; DB 21; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-148;  
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 VLSAFHHTLQLGPREQARNASCAGGPPADRRFRPPNLRSPWYRISYDPRYRL 67  
 Db 33 VLSAFHHTLQLGPREQARNASCAGGPPADRRFRPPNLRSPWYRISYDPRYRL 92  
 QY 68 PEAYCLRCGLTGLFGEDVRFSPVYMTVLRTPACAGGRSVYTEAYVTIPVGCCTC 127  
 Db 93 PEAYCLRCGLTGLFGEDVRFSPVYMTVLRTPACAGGRSVYTEAYVTIPVGCCTC 152  
 QY 128 VPEPEKDADSIINSIDKQAKLLIGPNDAPAGP 160  
 Db 153 VPEPEKDADSIINSIDKQAKLLIGPNDAPAGP 185  
 RESULT 4  
 AAY70655  
 ID AAY70655 standard; Protein; 186 AA.



```

XX DT 18-JUL-2000 (first entry)
XX DE Mature human Ztgbeta-9 variant protein.
XX KW Human transforming growth factor beta-9; Ztgb beta-9;
XX KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;
XX KW anyotrophic lateral sclerosis; ALS; Parkinson's disease;
XX KW peripheral neuropathy; demyelinating disease; multiple sclerosis;
XX KW antiviral; cytostatic.
XX OS Homo sapiens.
XX PN WO200015798-A2.
XX PD 23-MAR-2000.
XX PF 17-SEP-1999; 99WO-US21677.
XX PR 17-SEP-1998; 98US-0154817.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI Presnell SR, Taft DW, Foley KP;
XX DR WPI; 2000-271436/23.
XX DR N-PSDB; AA252201.
XX PT Polynucleotides encoding a novel transforming growth factor beta-9
XX PT polypeptide, designated Ztgb beta-9, useful as an antiviral and
XX PT antiproliferative agent
XX PS Claim 6; Page 93-94; 97pp; English.
XX CC The present sequence is a mature variant human transforming growth
XX CC factor beta-9 protein, designated Ztgb beta-9. Human Ztgb beta-9 was
XX CC isolated from an arrayed pituitary gland cDNA plasmid library by PCR
XX CC screening. This can be used to treat a variety of neurodegenerative
XX CC diseases such as anyotrophic lateral sclerosis (ALS), Alzheimer's
XX CC disease, Huntington's disease, Parkinson's disease and peripheral
XX CC neuropathies, or demyelinating diseases including multiple sclerosis.
XX CC Ztgb beta-9 peptides have antiviral activity and may also be used to
XX CC regulate the proliferation, differentiation and apoptosis of neurons,
XX CC glial cells, lymphocytes, hematopoietic cells and stromal cells.
XX SQ Sequence 187 AA;

Query Match 95.6%; Score 153; DB 21; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.2e-148;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 VLSAFHTLQLGPREQARNASCPAGGRPADRRFRPPPTNLRSPWAVRISYDPAARYPYL 67
DB 35 vlsafhtlqlgpreqarnascpaggrpadrrfrppptnlrsvspwayrisydypryl 94
QY 68 PEAYCLRCGLTGLFGEDVRFRSAPVYMPVTVLLRRTPCACGGRSVYTEAYVTIPVGCTC 127
DB 95 peayclrcglgtlfgedvrfrsapvymptvllrrtpacaggrsvyteayvtipvgctc 154
QY 128 VPEPEKDADTSINSDIKQAKLLIGPNDAPAGP 160
DB 155 vpepekdaidsinsidkqaklligpndapagp 187

RESULT 7
AAB07595
ID AAB07595 standard; Protein; 202 AA.
XX AC AAB07595;
XX AC AAB07595;
XX DT 07-NOV-2000 (first entry)
XX KW A human interleukin-173 polypeptide.

```

```

DE XX A human interleukin (IL) 173 polypeptide.
XX KW Interleukin; IL-17; CTLA-8; IL-170; IL-172; IL-173; IL-174; IL-176;
XX KW IL-177; IL-171; cell proliferation; cancer.
XX OS Homo sapiens.
XX Key Location/Qualifiers
XX Peptide 1..17
XX FT /note= "signal peptide"
XX FT Protein 18..202
XX FT /note= "mature protein"
XX PN WO200042188-A2.
XX PD 20-JUL-2000.
XX PF 10-JAN-2000; 2000WO-US000006.
XX PR 11-JAN-1999; 99US-0228822.
XX PA (SCHE ) SCHERING CORP.
XX PI Gorman DM, Bazan JF, Kastelein RA;
XX DR WPI; 2000-466130/40.
XX DR N-PSDB; AAA58983.
XX PT New isolated polynucleotide encoding a mammalian Interleukin-17 like
XX PT protein used to identify genes for homologous proteins -
XX PS Claim 11; Page 13-14; 111pp; English.
XX CC The present sequence represents an interleukin-173 (IL-173) polypeptide.
XX CC The polypeptide is an IL-17-like (CTLA-8 related) protein. It is a
XX CC member of a new group of interleukins, IL-170 polypeptides. The members
XX CC comprise IL-172, IL-173, IL-174, IL-176, IL-177, and IL-171. IL-170
XX CC protein can be used to treat abnormal proliferation e.g. cancer
XX CC or degenerative conditions. Antibodies can be used in diagnostic
XX CC methods to detect over production of IL-170 protein in cells or body
XX CC fluids.
XX SQ Sequence 202 AA;

Query Match 95.6%; Score 153; DB 21; Length 202;
Best Local Similarity 100.0%; Pred. No. 1.3e-148;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 VLSAFHTLQLGPREQARNASCPAGGRPADRRFRPPPTNLRSPWAVRISYDPAARYPYL 67
DB 50 vlsafhtlqlgpreqarnascpaggrpadrrfrppptnlrsvspwayrisydypryl 109
QY 68 PEAYCLRCGLTGLFGEDVRFRSAPVYMPVTVLLRRTPCACGGRSVYTEAYVTIPVGCTC 127
DB 110 peayclrcglgtlfgedvrfrsapvymptvllrrtpacaggrsvyteayvtipvgctc 169
QY 128 VPEPEKDADTSINSDIKQAKLLIGPNDAPAGP 160
DB 170 vpepekdaidsinsidkqaklligpndapagp 202

RESULT 8
AAB07689
ID AAB07689 standard; Protein; 202 AA.
XX AC AAB07689;
XX DT 07-NOV-2000 (first entry)
XX KW A human interleukin-173 polypeptide.
XX DE Interleukin; IL-171; cytokine; CTLA-8; IL-17; IL-175; IL-172; IL-173;
XX KW Interleukin; IL-171; cytokine; CTLA-8; IL-17; IL-175; IL-172; IL-173;

```

KW	IL-174; IL-176; IL-177; cell proliferation; cancer.
XX	Homo sapiens.
XX	
Key	Location/Qualifiers
Peptide	1..17 "signal peptide"
FT	/note=
FT	18..202 "mature protein"
FT	/note=
FT	50 "phosphorylation site"
FT	/note=
FT	50..53 "cAMP protein kinase phosphorylation site"
FT	/note=
FT	53 "phosphorylation site"
FT	/note=
FT	57..61 "myristoylation site"
FT	/note=
FT	57..61 "myristoylation site"
FT	/note=
FT	66..69 "CAMP protein kinase phosphorylation site"
FT	/note=
FT	66..69 "CAMP protein kinase phosphorylation site"
FT	/note=
FT	72..75 "CAMP protein kinase phosphorylation site"
FT	/note=
FT	72..75 "CAMP protein kinase phosphorylation site"
FT	/note=
FT	72 "CAMP protein kinase phosphorylation site"
FT	/note=
FT	"phosphorylation site"
FT	75
FT	/note=
FT	80 "phosphorylation site"
FT	/note=
FT	82 "phosphorylation site"
FT	/note=
FT	82..84 "phosphorylation site"
FT	/note=
FT	113 "calcium phosphorylation site"
FT	/note=
FT	"phosphorylation site"
FT	113..116
FT	/note=
FT	"CAMP protein kinase phosphorylation site"
FT	113..116
FT	/note=
FT	"CAMP protein kinase phosphorylation site"
FT	116
FT	/note=
FT	"phosphorylation site"
FT	164..166
FT	/note=
FT	"myristoylation site"
FT	164..166
FT	/note=
FT	"myristoylation site"
FT	166..168
FT	/note=
FT	"calcium phosphorylation site"
XX	
PX	WO200042187-A1.
PN	
XX	
PD	20-JUL-2000.
XX	
PX	10-JAN-2000; 2000WO-US000005.
XX	
PX	11-JAN-1999; 99US-0229402.
PR	
XX	(SCHE ) SCHERING CORP.
PA	Gorman DM, Bazan JF, Kastelein RA;
XX	
PI	
XX	
DR	WPI; 2000-476060/41.
DR	N-PSDB; AAA59155.
XX	
PT	New DNA sequence encoding a mammalian homolog of CTRIA-8, designated interleukin-171 (IL-171), useful for recombinant production of IL-171 which can be used for treating conditions associated with abnormal physiology or development -
PT	
XX	
PS	Disclosure; Page 16-17; lllpp; English.

XX		The present sequence represents an interleukin (IL)-173 polypeptide.
CC		It is a mammalian homologue of the cytokine designated CTLA-8 (also
CC		referred to as IL-17). The specification also describes homologues
CC		IL-171, IL-175, IL-172, IL-173, IL-174, IL-176 and IL-177. The DNA
CC		sequence encoding IL-171 is useful for identifying genes, mRNA and
CC		cDNA molecules which code for related or homologous proteins. The
CC		IL-171 protein, antibodies against IL-171, and compounds which have
CC		binding affinity to IL-171 are useful in treatment of conditions
CC		associated with abnormal physiology or development, including abnormal
CC		proliferation, e.g. cancerous conditions, or degenerative conditions.
CC		The IL-171 protein can be used in kits and assay methods for identifying
CC		compounds that selectively bind to IL-171.
XX		
SQ	Sequence	202 AA;
	Query Match	95.6%; Score 153; DB 21; Length 202;
	Best Local Similarity	100.0%; Pred. No. 1.3e-148;
	Matches 153; Conservative	0; Mismatches 0; Indels 0; Gaps
Qy	8	VLSAFHHTLQLGPREQARNASCAGRPADRRPPTNLRSVSPWAYRTSYDPARYPRYL 67 
Db	50	vlsafhtqlgprqearnascpagrpadrrfpptnlrsvspwayrtisydparypryl 109 
Qy	68	PEAYCLRCGLTGLFGBEDYRFSAPVYMTVVLRTPACAGRSYYTEAYVTIPVGCTC 127 
Db	110	peayclrcgltgfgedvrfisapvytmvtvlrrtpacagrsyyteayvtipvgctc 169 
Qy	128	VPEPEKDADSINSSIDKQGAKLLGPNDA PAGP 160 
Db	170	vpepek dads ins s i d k g g a k l l g p n d a pagp 202 
RESULT	9	
AAY70653		
ID	AAY70653 standard; Protein: 202 AA.	
XX		
AC	AAY70653;	
XX		
DT	18-JUL-2000 (first entry)	
XX		
DE	Human transforming growth factor beta-9, Ztgbeta-9 protein.	
XX		
KW	Human transforming growth factor beta-9; Ztgf beta-9;	
KW	Alzheimer's disease; neurodegenerative disease; Huntington's disease;	
KW	amyotrophic lateral sclerosis; ALS; Parkinson's disease;	
KW	peripheral neuropathy; demyelinating disease; multiple sclerosis;	
KW	antiviral; cytostatic.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..15
FT		/label= Signal_peptide
FT	Protein	16..202
FT		/note= "Mature transforming growth factor beta-9"
XX		
PN	WO200015798-A2.	
XX		
PD	23-MAR-2000.	
XX		
PF	17-SEP-1999; 99WO-US21677.	
XX		
PR	17-SEP-1998; 98US-0154817.	
XX		
PA	(ZYMO ) ZYMOGENETICS INC.	
XX		
PI	Presnell SR, Taft DW, Foley KP;	
XX		
DR	WPI; 2000-271436/23.	
XX	N-PSDB; AAZ52195.	
XX		

PT Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated ztgf beta-9, useful as an antiviral and  
PT antiproliferative agent  
PS Claim 6; Page 83; 97pp; English.  
XX  
CC The present sequence is the human transforming growth factor beta-9,  
CC designated ztgf beta-9. Human ztgf beta-9 was isolated from an arrayed  
CC pituitary gland cDNA plasmid library by PCR screening. This can be used  
CC to treat a variety of neurodegenerative diseases such as anyotrophic  
CC lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
CC Parkinson's disease and peripheral neuropathies, or demyelinating  
CC diseases including multiple sclerosis. Ztgf beta-9 peptides have  
CC antiviral activity and may also be used to regulate the proliferation,  
CC differentiation and apoptosis of neurons, glial cells, lymphocytes,  
CC hematopoietic cells and stromal cells.  
XX  
SQ Sequence 202 AA;  
  
Query Match 95.6%; Score 153; DB 21; Length 202;  
Best Local Similarity 100.0%; Pred. No. 1.3e-148;  
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 8 VLSAFHHTLQGPREQARNASCPAGRPADRRFRPPTNLRVSPWAYRISYDPAARYL 67  
Db 50 vlslafhtlqlgpreqarnascpagrpadrfrpptnlrvspwayrisydydaryl 109  
  
QY 68 PEAYCLRCGLTGLFGEDVRFRSAPVYMTVLLRTPACAGRSVYTEAYVTIPVGCTC 127  
Db 110 peayclrcglgtlfgedvrfrsapvymptvllrrtpacaggrsvyteayvtipvgctc 169  
  
QY 128 VPEPEKDADSINSSIDKQAKLLGPNADAPAG 160  
Db 170 vpepekadsinsidkqaklllgpnadapag 202  
  
RESULT 10  
AAU70662  
ID AAU70662 standard; Protein; 209 AA.  
XX  
AC AAU70662;  
XX  
DT 18-JUL-2000 (first entry)  
XX  
DE Human transforming growth factor beta-9, ztgf beta-9 variant protein.  
XX  
KW Human transforming growth factor beta-9; ztgf beta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW anyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytostatic.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..22  
FT /label= Signal\_peptide  
XX  
PN W0200015798-A2.  
XX  
PD 23-MAR-2000.  
XX  
PF 17-SEP-1999; 99WO-US21677.  
XX  
PR 17-SEP-1998; 98US-0154817.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Presnell SR, Taft DW, Foley KP;  
XX  
DR WPI; 2000-271436/23.  
DR N-PSDB; AA252201.

XX  
PT Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated ztgf beta-9, useful as an antiviral and  
PT antiproliferative agent  
PS Claim 6; Page 92-93; 97pp; English.  
XX  
CC The present sequence is a variant of human transforming growth  
CC factor beta-9, designated ztgf beta-9. Human ztgf beta-9 was isolated  
CC from an arrayed pituitary gland cDNA plasmid library by PCR screening.  
CC This can be used to treat a variety of neurodegenerative diseases such  
CC as anyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's  
CC disease, Parkinson's disease and peripheral neuropathies, or  
CC demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides  
CC have antiviral activity and may also be used to regulate the  
CC proliferation, differentiation and apoptosis of neurons, glial cells,  
CC lymphocytes, hematopoietic cells and stromal cells.  
XX  
SQ Sequence 209 AA;  
  
Query Match 95.6%; Score 153; DB 21; Length 209;  
Best Local Similarity 100.0%; Pred. No. 1.3e-148;  
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 8 VLSAFHHTLQGPREQARNASCPAGRPADRRFRPPTNLRVSPWAYRISYDPAARYL 67  
Db 57 vlslafhtlqlgpreqarnascpagrpadrfrpptnlrvspwayrisydydaryl 116  
  
QY 68 PEAYCLRCGLTGLFGEDVRFRSAPVYMTVLLRTPACAGRSVYTEAYVTIPVGCTC 127  
Db 117 peayclrcglgtlfgedvrfrsapvymptvllrrtpacaggrsvyteayvtipvgctc 176  
  
QY 128 VPEPEKDADSINSSIDKQAKLLGPNADAPAG 160  
Db 177 vpepekadsinsidkqaklllgpnadapag 209  
  
RESULT 11  
AAU04953  
ID AAU04953 standard; Protein; 202 AA.  
XX  
AC AAU04953;  
XX  
DT 24-OCT-2001 (first entry)  
XX  
DE Human Interleukin 17D ligand, IL-17D.  
XX  
KW Human; Interleukin-17D ligand; IL-17D; agonist; antagonist;  
KW PR021175; DNA 173894-2947; systemic lupus erythematosus;  
KW rheumatoid arthritis; osteoarthritis; diabetes mellitus;  
KW allergic disease; asthma; demyelinating disease;  
KW degenerative cartilaginous disorder; transplantation associated disease.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..15  
FT /label= Signal\_peptide  
FT Protein 16..202  
FT /label= Mature\_IL\_17D  
FT Region 17..23  
FT /note= "N-myristoylation site"  
FT Region 21..25  
FT /note= "Amidation site"  
FT Region 49..55  
FT /note= "N-myristoylation site"  
FT Modified-site 68..72  
FT /note= "Asn is glycosylated"  
FT Region 74..80  
FT /note= "N-myristoylation site"  
FT Region 97..106  
FT /note= "Tyrosine kinase phosphorylation site"



XX DE A rat interleukin-173 polypeptide.  
XX DE Interleukin; IL-171; cytokine; CTLA-8; IL-17; IL-175; IL-172; IL-173;  
KW IL-174; IL-176; IL-177; cell proliferation; cancer.  
XX OS Rattus sp.  
XX PN WO200042187-A1.  
XX PD 20-JUL-2000.  
XX PF 10-JAN-2000; 2000WO-US00005.  
XX PR 11-JAN-1999; 99US-0229402.  
XX PA (SCHE ) SCHERING CORP.  
XX PI Gorman DM, Bazan JF, Kastelein RA;  
XX DR WPI; 2000-476060/41.  
XX DR N-PSDB; AAA59156.  
XX PT New DNA sequence encoding a mammalian homolog of CTLA-8, designated  
PT interleukin-171 (IL-171), useful for recombinant production of IL-171  
PT which can be used for treating conditions associated with abnormal  
PT physiology or development -  
XX PS Disclosure; Page 97; 11pp; English.  
XX CC The present sequence represents an interleukin (IL)-173 polypeptide.  
CC It is a mammalian homologue of the cytokine designated CTLA-8 (also  
CC referred to as IL-17). The specification also describes homologues  
CC IL-171, IL-175, IL-172, IL-173, IL-174, IL-176, and IL-177. The DNA  
CC sequence encoding IL-171 is useful for identifying genes, mRNA and  
CC cDNA molecules which code for related or homologous proteins. The  
CC IL-171 protein, antibodies against IL-171, and compounds which have  
CC binding affinity to IL-171 are useful in treatment of conditions  
CC associated with abnormal physiology or development, including abnormal  
CC proliferation, e.g. cancerous conditions, or degenerative conditions.  
CC The IL-171 protein can be used in kits and assay methods for identifying  
CC compounds that selectively bind to IL-171.  
XX SQ Sequence 151 AA;  
  
Query Match 60.6%; Score 97; DB 21; Length 151;  
Best Local Similarity 100.0%; Pred. No. 2.1e-91;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 43 PTNLSRSPWAYRISYDPARYRPLPEAYCLCRGCLTGLFGEDVFRFSAPVYMTWVLR 102  
Db 55 ptnlsrspwayrisydparyrplpeayclrgcltglfgeedvfrfsapvymptwvlr 114  
  
QY 103 RTPACAGGRSVYTEAYVTIPVGTCTVPEPEKADSDIN 139  
Db 115 rtpacaggrsvyteayvtipvgctcvpepekadsin 151  
  
RESULT 14  
AAY70666  
ID AAY70666 standard; Protein; 57 AA.  
XX AC AAY70666;  
XX DT 18-JUL-2000 (first entry)  
XX DE Human transforming growth factor beta-9, ztgbeta-9 peptide-6.  
XX KW Human transforming growth factor beta-9; Ztgbeta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
  
A rat interleukin-173 polypeptide.  
Homo sapiens.  
WO200015798-A2.  
23-MAR-2000.  
17-SEP-1999; 99WO-US21677.  
17-SEP-1998; 98US-0154817.  
(ZYMO ) ZYMOGENETICS INC.  
Presnell SR, Taft DW, Foley KP;  
WPI; 2000-271436/23.  
Polynucleotides encoding a novel transforming growth factor beta-9  
polypeptide, designated Ztgbeta-9, useful as an antiviral and  
antiproliferative agent -  
Claim 6; Page 94-95; 97pp; English.  
The present sequence is an epitope-bearing peptide derived from the human  
transforming growth factor beta-9, designated Ztgbeta-9. These peptides  
are used to raise antibodies, including anti-idiotypic antibodies. This  
can be used to treat a variety of neurodegenerative diseases such as  
amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's  
disease, Parkinson's disease and peripheral neuropathies, or  
demyelinating diseases including multiple sclerosis. Ztgbeta-9 peptides  
have antiviral activity and may also be used to regulate the  
proliferation, differentiation and apoptosis of neurons, glial cells,  
lymphocytes, hematopoietic cells and stromal cells.  
Sequence 57 AA;  
  
Query Match 35.6%; Score 57; DB 21; Length 57;  
Best Local Similarity 100.0%; Pred. No. 7.9e-51;  
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 102 RRTPCAGGRSVYTEAYVTIPVGTCTVPEPEKDADSINSIDKOGAKLLGPNDA 158  
Db 1 rrtpacaggrsvyteayvtipvgctcvpepekadsinsidkqgakllgpnadapa 57  
  
RESULT 15  
AAY70664  
ID AAY70664 standard; Protein; 54 AA.  
XX AC AAY70664;  
XX DT 18-JUL-2000 (first entry)  
XX DE Human transforming growth factor beta-9, Ztgbeta-9 peptide-4.  
XX KW Human transforming growth factor beta-9; Ztgbeta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytostatic.  
XX OS Homo sapiens.  
XX PN WO200015798-A2.  
XX PD 23-MAR-2000.  
XX PF 17-SEP-1999; 99WO-US21677.  
XX PR 17-SEP-1998; 98US-0154817.  
XX SQ

PA (ZYMO ) ZYMOGENETICS INC.  
XX Presnell SR, Taft DW, Foley KP;  
XX WPI; 2000-271436/23.  
XX Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated ztgf beta-9, useful as an antiviral and  
PT antiproliferative agent -  
XX Claim 6; Page 94; 97pp; English.  
XX The present sequence is an epitope-bearing peptide derived from the human  
CC transforming growth factor beta-9, designated ztgf beta-9. These peptides  
CC are used to raise antibodies, including anti-idiotypic antibodies. This  
CC can be used to treat a variety of neurodegenerative diseases such as  
CC amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's  
CC disease, Parkinson's disease and peripheral neuropathies, or  
CC demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides  
CC have antiviral activity and may also be used to regulate the  
CC proliferation, differentiation and apoptosis of neurons, glial cells,  
CC lymphocytes, hematopoietic cells and stromal cells.  
XX Sequence 54 AA;  
QY 17 QLGPREARNASCAGGRPADRRFRPTNLRVSPWAYRISYDPARYLP 70  
DB 1 qlgprearnascaggrpadrrfrptnlrsvspwayrissydparylp 54  
RESULT 16  
AAB07688  
ID AAB07688 standard; Protein; 103 AA.  
XX AAB07688;  
XX 07-NOV-2000 (first entry)  
DE A human interleukin-173 polypeptide.  
XX Interleukin; IL-171; cytokine; CTLA-8; IL-17; IL-175; IL-172; IL-173;  
KW IL-174; IL-176; IL-177; cell proliferation; cancer.  
XX Homo sapiens.  
OS WO200042187-A1.  
PN 20-JUL-2000.  
XX 10-JAN-2000; 2000WO-US000005.  
XX 11-JAN-1999; 99US-0229402.  
XX (SCHE ) SCHERING CORP.  
PI Gorman DM, Bazan JF, Kastelein RA;  
XX WPI; 2000-476060/41.  
DR N-PSDB; AAA59154.  
XX New DNA sequence encoding a mammalian homolog of CTLA-8, designated  
PT interleukin-171 (IL-171), useful for recombinant production of IL-171  
PT which can be used for treating conditions associated with abnormal  
PT physiology or development -  
XX Disclosure; Page 15; 111pp; English.  
XX The present sequence represents an interleukin (IL)-173 polypeptide.

CC It is a mammalian homologue of the cytokine designated CTLA-8 (also  
CC referred to as IL-17). The specification also describes homologues  
CC IL-171, IL-175, IL-172, IL-173, IL-174, IL-176, and IL-177. The DNA  
CC sequence encoding IL-171 is useful for identifying genes, mRNA and  
CC cDNA molecules which code for related or homologous proteins. The  
CC IL-171 protein, antibodies against IL-171, and compounds which have  
CC binding affinity to IL-171 are useful in treatment of conditions  
CC associated with abnormal physiology or development, including abnormal  
CC proliferation, e.g. cancerous conditions, or degenerative conditions.  
CC The IL-171 protein can be used in kits and assay methods for identifying  
CC compounds that selectively bind to IL-171.  
XX Sequence 103 AA;  
QY 43 PTNLRVSPWAYRISYDPARYLPAYCLRCGCLTGLFGEDVRF 91  
DB 55 ptnlrsvspwayrissydparylpayclrcgcltglfgesdvrf 103  
RESULT 17  
AAY70661  
ID AAY70661 standard; Protein; 34 AA.  
XX AAY70661;  
XX 18-JUL-2000 (first entry)  
DE Human transforming growth factor beta-9, ztgf beta-9 peptide-3.  
XX Human transforming growth factor beta-9; ztgf beta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytostatic.  
XX Homo sapiens.  
OS WO200015798-A2.  
PN 23-MAR-2000.  
XX 17-SEP-1999; 99WO-US21677.  
XX 17-SEP-1998; 98US-0154817.  
XX (ZYMO ) ZYMOGENETICS INC.  
PI Presnell SR, Taft DW, Foley KP;  
XX WPI; 2000-271436/23.  
XX Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated ztgf beta-9, useful as an antiviral and  
PT antiproliferative agent -  
XX Claim 6; Page 90; 97pp; English.  
XX The present sequence is an epitope-bearing peptide derived from the human  
CC transforming growth factor beta-9, designated ztgf beta-9. These peptides  
CC are used to raise antibodies, including anti-idiotypic antibodies. This  
CC can be used to treat a variety of neurodegenerative diseases such as  
CC amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's  
CC disease, Parkinson's disease and peripheral neuropathies, or  
CC demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides  
CC have antiviral activity and may also be used to regulate the  
CC proliferation, differentiation and apoptosis of neurons, glial cells,  
CC lymphocytes, hematopoietic cells and stromal cells.  
XX Query Match 30.6%; Score 49; DB 21; Length 103;  
Best Local Similarity 100.0%; Pred. No. 2.1e-42;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 43 PTNLRVSPWAYRISYDPARYLPAYCLRCGCLTGLFGEDVRF 91  
DB 55 ptnlrsvspwayrissydparylpayclrcgcltglfgesdvrf 103

SQ Sequence 34 AA;

Query Match 21.2%; Score 34; DB 21; Length 34;  
Best Local Similarity 100.0%; Pred. No. 1.8e-27;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 CVPEPKDADINSIDKQAKLLGPNAPAGP 160  
|||||  
Db 1 cvpekdadsinsidkqgakkllgpnadapag 34

RESULT 18

AAAY70658  
ID AAY70658 standard; Protein; 183 AA.

XX AC AAY70658;

XX XT 18-JUL-2000 (first entry)

XX DE Mature murine transforming growth factor beta-9, Ztgfbeta-9 protein.

XX KW Murine transforming growth factor beta-9; Ztgf beta-9;

KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;

KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;

KW peripheral neuropathy; demyelinating disease; multiple sclerosis;

KW antiviral; cytostatic.

XX OS Mus sp.

XX PN WO200015798-A2.

XX PD 23-MAR-2000.

XX PF 17-SEP-1999; 99WO-US21677.

XX PR 17-SEP-1998; 98US-0154817.

XX PA (ZYMO ) ZYMOGENETICS INC.

XX PI Presnell SR, Taft DW, Foley KP;

XX DR WPI; 2000-271436/23.

DR N-PSDB; AA252198.

XX PT Polynucleotides encoding a novel transforming growth factor beta-9  
polypeptide, designated ztgf beta-9, useful as an antiviral and  
antiproliferative agent

XX PS Claim 6; Page 89; 97pp; English.

XX CC The present sequence is the mature murine transforming growth factor  
beta-9, designated Ztgf beta-9. The signal sequence extends from  
amino acid residues 1 to 22. Murine Ztgf beta-9 was found to be highly  
expressed in the HCL hypothalamic cell line. This can be used  
to treat a variety of neurodegenerative diseases such as amyotrophic  
lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
Parkinson's disease and peripheral neuropathies, or demyelinating  
diseases including multiple sclerosis. Ztgf beta-9 peptides have  
antiviral activity and may also be used to regulate the proliferation,  
differentiation and apoptosis of neurons, glial cells, lymphocytes,  
hematopoietic cells and stromal cells.

XX SQ Sequence 183 AA;

Query Match 16.9%; Score 27; DB 21; Length 183;  
Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 VLSAFHHTLQLGPREQARNASCPAGGR 34

|||||  
Db 35 vlsafhtlqlgpreqarnascpaggr 61

RESULT 19

AAAB07597  
ID AAB07597 standard; Protein; 205 AA.

XX AC AAB07597;

XX DT 07-NOV-2000 (first entry)

XX DE A murine interleukin (IL) 173 polypeptide.

XX KW Interleukin; IL-17; CTLA-8; IL-170; IL-172; IL-173; IL-174; IL-176;  
KW IL-177; IL-171; cell proliferation; cancer.

XX OS Mus sp.

XX FH Key Location/Qualifiers  
FT Peptide 1..24  
FT Protein /note= "signal peptide"

FT Protein 25..205 /note= "mature protein"

FT Modified-site 50

FT Modified-site /note= "phosphorylation site"

FT Modified-site 51..53

FT Modified-site /note= "N-glycosylation site"

FT Modified-site 53

FT Modified-site /note= "phosphorylation site"

FT Modified-site 57..61

FT Modified-site /note= "myristoylation site"

FT Modified-site 80

FT Modified-site /note= "phosphorylation site"

FT Modified-site 82

FT Modified-site /note= "phosphorylation site"

FT Modified-site 82..84

FT Modified-site /note= "calcium phosphorylation site"

FT Modified-site 101..105

FT Modified-site /note= "myristoylation site"

FT Modified-site 113

FT Modified-site /note= "phosphorylation site"

FT Modified-site 116

FT Modified-site /note= "phosphorylation site"

FT Modified-site 164..166

FT Modified-site /note= "N-glycosylation site"

FT Modified-site 166..168

FT Modified-site /note= "calcium phosphorylation site"

XX PN WO200042188-A2.

XX PD 20-JUL-2000.

XX PF 10-JAN-2000; 2000WO-US000006.

XX PR 11-JAN-1999; 99US-0228822.

XX PA (SCHE ) SCHERING CORP.

XX PI Gorman DM, Bazan JF, Kastelein RA;

XX DR WPI; 2000-466130/40.

DR N-PSDB; AAA58985.

XX PT New isolated polynucleotide encoding a mammalian Interleukin-17 like  
protein used to identify genes for homologous proteins

XX PS Claim 11; Page 15-16; 111pp; English.

XX CC The present sequence represents an interleukin-173 (IL-173) polypeptide.  
The polypeptide is an IL-17-like (CTLA-8 related) protein. It is a  
member of a new group of interleukins, IL-170 polypeptides. The members  
comprise IL-172, IL-173, IL-174, IL-176, IL-177, and IL-171. IL-170  
protein can be used to treat abnormal proliferation e.g. cancer  
or degenerative conditions. Antibodies can be used in diagnostic

CC methods to detect over production of IL-170 protein in cells or body  
XX fluids.  
SQ Sequence 205 AA;

Query Match 16.9%; Score 27; DB 21; Length 205;  
Best Local Similarity 100.0%; Pred. No. 1.3e-19;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VLSAFHHTLQLGPREQARNASCPAGGR 34  
|||||  
Db 57 vlsafhhtlqlgpreqarnascpaggr 83

RESULT 20  
AAB07691  
ID AAB07691 standard; Protein; 205 AA.

AC AAB07691;

DT 07-NOV-2000 (first entry)

DE A murine interleukin-173 polypeptide.

XX Interleukin; IL-171; cytokine; CTIA-8; IL-17; IL-175; IL-172; IL-173;  
KW IL-174; IL-176; IL-177; cell proliferation; cancer; ss.

OS Mus sp.

FH Key Location/Qualifiers

FT Peptide 1..24

FT /note= "signal peptide"

FT Modified-site 4..6

FT Protein /note= "protein kinase C phosphorylation site"

FT /note= "mature protein"

FT Modified-site 50..53

FT /note= "cAMP protein kinase site"

FT Modified-site 50

FT /note= "phosphorylation site"

FT Modified-site 53

FT /note= "phosphorylation site"

FT Modified-site 51..53

FT /note= "N-glycosylation site"

FT Modified-site 57..61

FT /note= "myristoylation site"

FT Modified-site 66..69

FT /note= "cAMP protein kinase site"

FT Modified-site 72..75

FT /note= "cAMP protein kinase site"

FT Modified-site 72

FT /note= "phosphorylation site"

FT Modified-site 75

FT /note= "phosphorylation site"

FT Modified-site 80

FT /note= "phosphorylation site"

FT Modified-site 82

FT /note= "phosphorylation site"

FT Modified-site 82..84

FT /note= "calcium phosphorylation site"

FT Modified-site 101..105

FT /note= "myristoylation site"

FT Modified-site 113

FT /note= "phosphorylation site"

FT Modified-site 113..116

FT /note= "cAMP protein kinase site"

FT Modified-site 116

FT /note= "phosphorylation site"

FT Modified-site 159..161

FT /note= "calcium phosphorylation site"

FT Modified-site 164..166

FT /note= "N-glycosylation site"

FT Modified-site 166..168  
FT /note= "calcium phosphorylation site"

PN WO200042187-A1.

XX 20-JUL-2000.

XX 10-JAN-2000; 2000WO-US000005.

XX 11-JAN-1999; 99US-0229402.

XX (SCHE ) SCHERING CORP.

XX Gorman DM, Bazan JF, Kastelein RA;

XX WPI; 2000-476060/41.

DR N-PSDB; AAA59157.

XX New DNA sequence encoding a mammalian homolog of CTIA-8, designated  
PT interleukin-171 (IL-171), useful for recombinant production of IL-171  
PT which can be used for treating conditions associated with abnormal  
PT physiology or development -

XX Disclosure; Page 17-18; 111pp; English.

XX The present sequence represents an interleukin (IL)-173 polypeptide.  
CC It is a mammalian homolog of the cytokine designated CTIA-8 (also  
CC referred to as IL-17). The specification also describes homologues  
CC IL-171, IL-175, IL-172, IL-173, IL-174, IL-176, and IL-177. The DNA  
CC sequence encoding IL-171 is useful for identifying genes, mRNA and  
CC cDNA molecules which code for related or homologous proteins. The  
CC IL-171 protein, antibodies against IL-171, and compounds which have  
CC binding affinity to IL-171 are useful in treatment of conditions  
CC associated with abnormal physiology or development, including abnormal  
CC proliferation, e.g. cancerous conditions, or degenerative conditions.  
CC The IL-171 protein can be used in kits and assay methods for identifying  
CC compounds that selectively bind to IL-171.

SQ Sequence 205 AA;

Query Match 16.9%; Score 27; DB 21; Length 205;

Best Local Similarity 100.0%; Pred. No. 1.3e-19;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VLSAFHHTLQLGPREQARNASCPAGGR 34  
|||||

Db 57 vlsafhhtlqlgpreqarnascpaggr 83  
|||||

RESULT 21

AAY70657

ID AAY70657 standard; Protein; 205 AA.

XX AAY70657;

XX 18-JUL-2000 (first entry)

DT Murine transforming growth factor beta-9, Ztgf beta-9 protein.

DE Murine transforming growth factor beta-9; Ztgf beta-9;

XX Alzheimer's disease; neurodegenerative disease; Huntington's disease;

XX amyotrophic lateral sclerosis; ALS; Parkinson's disease;

XX peripheral neuropathy; demyelinating disease; multiple sclerosis;

XX antiviral; cytostatic.

XX Mus sp.

XX Location/Qualifiers

FT Key 1..22

FT Peptide /label= Signal\_peptide

XX WO200015798-A2.

XX PD 23-MAR-2000.  
 XX XX  
 XX PF 17-SEP-1999; 99WO-US21677.  
 XX XX  
 XX PR 17-SEP-1998; 98US-0154817.  
 XX XX  
 XX PA (ZYMO) ZYMOGENETICS INC.  
 XX XX  
 XX PI Presnell SR, Taft DW, Foley KP;  
 XX XX  
 XX DR WPI: 2000-271436/23.  
 XX DR N-PSDB; AA521198.  
 XX XX  
 XX PT Polynucleotides encoding a novel transforming growth factor beta-9  
 XX PT polypeptide, designated ztgf beta-9, useful as an antiviral and  
 XX PT antiproliferative agent.  
 XX PS Claim 6; Page 88; 97pp; English.  
 XX XX  
 XX CC The present sequence encodes murine transforming growth factor beta-9,  
 XX CC designated ztgf beta-9. Murine ztgf beta-9 was found to be highly  
 XX CC expressed in the HCL hypothalamic cell line. This can be used  
 XX CC to treat a variety of neurodegenerative diseases such as amyotrophic  
 XX CC lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
 XX CC Parkinson's disease and peripheral neuropathies, or demyelinating  
 XX CC diseases including multiple sclerosis. Ztgf beta-9 peptides have  
 XX CC antiviral activity and may also be used to regulate the proliferation,  
 XX CC differentiation and apoptosis of neurons, glial cells, lymphocytes,  
 XX CC hematopoietic cells and stromal cells.  
 XX SQ Sequence 205 AA;  
 XX XX  
 XX Query Match 16.9%; Score 27; DB 21; Length 205;  
 XX Best Local Similarity 100.0%; Pred. No. 1.3e-19;  
 XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX XX  
 XX QY 8 VLSAFHHTLQLGPREQARNASCPAGGR 34  
 XX DB 57 vlsafhhtlqlgpreqarnascpaggr 83  
 XX XX  
 XX RESULT 22  
 XX AAE05341  
 XX ID AAE05341 standard; Protein; 205 AA.  
 XX XX  
 XX AC AAE05341;  
 XX XX  
 XX DT 12-SEP-2001 (first entry)  
 XX XX  
 XX DE Mouse zcyto7 protein.  
 XX XX  
 XX KW Mouse; cytostatic; antiinflammatory; immunoregulatory; tissue integrity;  
 XX KW wound healing; immune response; vaccine; cancer; asthma; allergy;  
 XX KW cell trafficking; therapy; zcyto7 protein; secreted protein; IL-17;  
 XX KW interleukin-17 related cytokine.  
 XX XX  
 XX OS Mus sp.  
 XX XX  
 XX PN WO200148192-A1.  
 XX XX  
 XX PD 05-JUL-2001.  
 XX XX  
 XX PF 21-DEC-2000; 2000WO-NZ00256.  
 XX XX  
 XX PR 23-DEC-1999; 99US-0171678.  
 XX PR 28-NOV-2000; 2000US-0724864.  
 XX XX  
 XX PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX XX  
 XX PI Watson JD, Murison JG;  
 XX XX

DR WPI: 2001-425665/45.  
 DR N-PSDB; AAD10108.  
 XX XX  
 XX PT Novel isolated polypeptide useful to isolate corresponding interacting  
 XX PT proteins or other compounds, to quantitatively determine levels of  
 XX PT interacting proteins or other compounds, and as therapeutic target -  
 XX PS Claim 6; Page 71-72; 101pp; English.  
 XX XX  
 XX CC The patent discloses novel polynucleotides and their corresponding  
 XX CC proteins which play a major role in induction of growth, cell migration  
 XX CC and proliferation, cell-cell interaction and the differentiation of  
 XX CC tissue-specific cells. These proteins are important in the maintenance  
 XX CC of tissue integrity and thus are important in wound healing. They are  
 XX CC useful in various assays to determine the biological activity, to raise  
 XX CC antibodies, to isolate corresponding interacting proteins or other  
 XX CC compounds, to quantitatively determine levels of interacting proteins or  
 XX CC other compounds, and as therapeutic target in a whole range of disease  
 XX CC states. Compositions comprising the novel proteins of the invention are  
 XX CC useful for treating mammalian disorders. Polynucleotides of the invention  
 XX CC are useful in genome and physical mapping, in positional cloning of  
 XX CC genes, to tag or identify an organism or its reproductive material (as  
 XX CC non-disruptive tags for marking organisms), and for the diagnosis and  
 XX CC treatment of mammalian diseases which is the consequence of inappropriate  
 XX CC expression of kinase genes. They are useful for promoting immune response  
 XX CC as part of a vaccine or anti-cancer treatment, as target for cancer  
 XX CC treatment, as immunoregulatory and anti-inflammatory molecule, as  
 XX CC diagnostic for specific types of cancer and for development of an  
 XX CC anti-cancer treatment, and as a target for antagonists in the treatment  
 XX CC of diseases such as asthma and allergy. They are also useful to inhibit  
 XX CC or enhance the activity of the soluble molecule that binds proteins of  
 XX CC the invention, for tissue and neural regeneration, to promote or block  
 XX CC cell trafficking, and as anti-inflammatory and/or vaccine adjuvant.  
 XX CC The present sequence is zcyto7 protein, a secreted protein from mouse  
 XX CC which belongs to the family of interleukin-17 (IL-17)-related cytokines.  
 XX SQ Sequence 205 AA;  
 XX XX  
 XX Query Match 16.9%; Score 27; DB 22; Length 205;  
 XX Best Local Similarity 100.0%; Pred. No. 1.3e-19;  
 XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX XX  
 XX QY 8 VLSAFHHTLQLGPREQARNASCPAGGR 34  
 XX DB 57 vlsafhhtlqlgpreqarnascpaggr 83  
 XX XX  
 XX RESULT 23  
 XX AAY70659  
 XX ID AAY70659 standard; Protein; 31 AA.  
 XX XX  
 XX AC AAY70659;  
 XX XX  
 XX DT 18-JUL-2000 (first entry)  
 XX XX  
 XX DE Human transforming growth factor beta-9, ztgbeta-9 peptide-1.  
 XX XX  
 XX KW Human transforming growth factor beta-9; ztgf beta-9;  
 XX KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
 XX KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
 XX KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
 XX KW antiviral; cytostatic.  
 XX XX  
 XX OS Homo sapiens.  
 XX XX  
 XX PN WO200015798-A2.  
 XX XX  
 XX PD 23-MAR-2000.  
 XX XX  
 XX PF 17-SEP-1999; 99WO-US21677.  
 XX XX  
 XX PR 17-SEP-1998; 98US-0154817.  
 XX XX

```
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI Presnell SR, Taft DW, Foley KP;
XX PI WPI; 2000-271436/23.
XX DR
XX PT Polynucleotides encoding a novel transforming growth factor beta-9
XX PT polypeptide, designated ztgf beta-9, useful as an antiviral and
XX PT antiproliferative agent
XX PS
XX PS Claim 6; Page 90; 97pp; English.
XX CC The present sequence is an epitope-bearing peptide derived from the human
XX CC transforming growth factor beta-9, designated ztgf beta-9. These peptides
XX CC are used to raise antibodies, including anti-idiotypic antibodies. This
XX CC can be used to treat a variety of neurodegenerative diseases such as
XX CC anyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's
XX CC disease, Parkinson's disease and peripheral neuropathies, or
XX CC demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides
XX CC have antiviral activity and may also be used to regulate the
XX CC proliferation, differentiation and apoptosis of neurons, glial cells,
XX CC lymphocytes, hematopoietic cells and stromal cells.
XX SQ Sequence 31 AA;

Query Match 14.4%; Score 23; DB 21; Length 31;
Best Local Similarity 100.0%; Pred. No. 3e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QLGPREQARNASCPAGGRPADRR 39
Db 1 qlgpreqarnascpaggrpadrr 23

RESULT 24
AAY70660
ID AAY70660 standard; Protein; 21 AA.
XX AC
XX AC AAY70660;
XX DT
XX DT 18-JUL-2000 (first entry)
XX DE Human transforming growth factor beta-9, ztgf beta-9 peptide-2.
XX KW Human transforming growth factor beta-9; ztgf beta-9;
XX KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;
XX KW anyotrophic lateral sclerosis; ALS; Parkinson's disease;
XX KW peripheral neuropathy; demyelinating disease; multiple sclerosis;
XX KW antiviral; cytostatic.
XX OS Homo sapiens.
XX PN WO200015798-A2.
XX PD 23-MAR-2000.
XX PF 17-SEP-1999; 99WO-US21677.
XX PR 17-SEP-1998; 98US-0154817.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI Presnell SR, Taft DW, Foley KP;
XX PI WPI; 2000-271436/23.
XX PT Polynucleotides encoding a novel transforming growth factor beta-9
XX PT polypeptide, designated ztgf beta-9, useful as an antiviral and
XX PT antiproliferative agent
XX PS
XX PS Claim 6; Page 90; 97pp; English.
```

```
XX CC The present sequence is an epitope-bearing peptide derived from the human
XX CC transforming growth factor beta-9, designated ztgf beta-9. These peptides
XX CC are used to raise antibodies, including anti-idiotypic antibodies. This
XX CC can be used to treat a variety of neurodegenerative diseases such as
XX CC anyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's
XX CC disease, Parkinson's disease and peripheral neuropathies, or
XX CC demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides
XX CC have antiviral activity and may also be used to regulate the
XX CC proliferation, differentiation and apoptosis of neurons, glial cells,
XX CC lymphocytes, hematopoietic cells and stromal cells.
XX SQ Sequence 21 AA;

Query Match 12.5%; Score 20; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.5e-13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 GEEDVRFRRSAPVYMPVTWLR 102
Db 1 geedvrfrrsapvympvtwlr 20

RESULT 25
AAB07596
ID AAB07596 standard; Protein; 44 AA.
XX AC
XX AC AAB07596;
XX DT 07-NOV-2000 (first entry)
XX DE A rat interleukin (IL) 173 polypeptide.
XX KW Interleukin; IL-17; CTLA-8; IL-170; IL-172; IL-173; IL-174; IL-176;
XX KW IL-177; IL-171; cell proliferation; cancer.
XX OS Rattus sp.
XX PN WO200042188-A2.
XX PD 20-JUL-2000.
XX PF 10-JAN-2000; 2000WO-US00006.
XX PR 11-JAN-1999; 99US-0228822.
XX PA (SCHE ) SCHERING CORP.
XX PI Gorman DM, Bazan JF, Kastelein RA;
XX PI WPI; 2000-466130/40.
XX DR N-PSDB; AAA58984.
XX DR
XX PT New isolated polynucleotide encoding a mammalian Interleukin-17 like
XX PT protein used to identify genes for homologous proteins -
XX PS Claim 11; Page 14; 111pp; English.
XX CC The present sequence represent an interleukin-173 (IL-173) polypeptide.
XX CC The polypeptide is an IL-17-like (CTLA-8 related) protein. It is a
XX CC member of a new group of interleukins, IL-170 polypeptides. The members
XX CC comprise IL-172, IL-173, IL-174, IL-176, IL-177, and IL-171. IL-170
XX CC protein can be used to treat abnormal proliferation e.g. cancer
XX CC or degenerative conditions. Antibodies can be used in diagnostic
XX CC methods to detect over production of IL-170 protein in cells or body
XX CC fluids.
XX SQ Sequence 44 AA;

Query Match 11.2%; Score 18; DB 21; Length 44;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
```

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 PRLPEAYCLRGGLTGL 81  
|||||  
Db 2 prylpeaycicrgcgtgl 19

RESULT 26  
AAV70665  
ID AAY70665 standard; Protein; 16 AA.  
XX AC AAY70665;  
XX DT 18-JUL-2000 (first entry)  
XX DE Human transforming growth factor beta-9, Ztgfbeta-9 peptide-5.  
XX KW Human transforming growth factor beta-9; Ztgfbeta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytostatic.  
XX OS Homo sapiens.  
XX PN WO200015798-A2.  
XX PD 23-MAR-2000.  
XX PF 17-SEP-1999; 99WO-US21677.  
XX KW Human transforming growth factor beta-9; Ztgfbeta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytostatic.  
XX PA (ZYMO ) ZYMOGENETICS INC.  
XX PI Presnell SR, Taft DW, Foley KP;  
XX WPI; 2000-271436/23.  
XX PD 23-MAR-2000.  
XX PF 17-SEP-1999; 99WO-US21677.  
XX PR 17-SEP-1998; 98US-0154817.  
XX PS Claim 6; Page 94; 97pp; English.  
XX CC The present sequence is an epitope-bearing peptide derived from the human transforming growth factor beta-9, designated Ztgfbeta-9. These peptides are used to raise antibodies, including anti-idiotypic antibodies. This can be used to treat a variety of neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease, Parkinson's disease and peripheral neuropathies, or demyelinating diseases including multiple sclerosis. Ztgfbeta-9 peptides have antiviral activity and may also be used to regulate the proliferation, differentiation and apoptosis of neurons, glial cells, lymphocytes, hematopoietic cells and stromal cells.  
XX SQ Sequence 16 AA;

Query Match 10.0%; Score 16; DB 21; Length 16;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 RPADRRFPPTNLRSV 49  
|||||  
Db 1 rpadrfrpbtnlrsv 16

RESULT 27  
AAV70667  
ID AAY70667 standard; Protein; 14 AA.  
XX AC AAY70667;

XX DT 18-JUL-2000 (first entry)  
XX DE Human transforming growth factor beta-9, Ztgfbeta-9 peptide-7.  
XX KW Human transforming growth factor beta-9; Ztgfbeta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytostatic.  
XX OS Homo sapiens.  
XX PN WO200015798-A2.  
XX PD 23-MAR-2000.  
XX PF 17-SEP-1999; 99WO-US21677.  
XX PR 17-SEP-1998; 98US-0154817.  
XX PA (ZYMO ) ZYMOGENETICS INC.  
XX PI Presnell SR, Taft DW, Foley KP;  
XX WPI; 2000-271436/23.  
XX PD Polynucleotides encoding a novel transforming growth factor beta-9 polypeptide, designated Ztgfbeta-9, useful as an antiviral and antiproliferative agent -  
XX PT antiproliferative agent -  
XX PS Claim 6; Page 95; 97pp; English.  
XX CC The present sequence is an epitope-bearing peptide derived from the human transforming growth factor beta-9, designated Ztgfbeta-9. These peptides are used to raise antibodies, including anti-idiotypic antibodies. This can be used to treat a variety of neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease, Parkinson's disease and peripheral neuropathies, or demyelinating diseases including multiple sclerosis. Ztgfbeta-9 peptides have antiviral activity and may also be used to regulate the proliferation, differentiation and apoptosis of neurons, glial cells, lymphocytes, hematopoietic cells and stromal cells.  
XX SQ Sequence 14 AA;

Query Match 8.8%; Score 14; DB 21; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.4e-07;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 SYDPARYPRVILPEA 70  
|||||  
Db 1 sydparyprylpea 14

RESULT 28  
AAV84562  
ID AAY84562 standard; Protein; 805 AA.  
XX AC AAY84562;  
XX DT 25-JUL-2000 (first entry)  
XX DE A human angiotensin converting enzyme-2 (ACE-2) protein.  
XX KW Human; angiotensin converting enzyme-2; ACE-2; angiotensin I; Ang.(1-9);  
KW blood pressure; hypertension; congestive heart failure; atherosclerosis;  
KW chronic heart failure; acute heart failure; myocardial infarction;  
KW renal failure.  
XX OS Homo sapiens.  
XX

FH Key Location/Qualifiers  
FT Peptide 1..18  
FT /note= "signal sequence"  
FT Domain 19..740  
FT /note= "extracellular domain"  
FT Domain 374..378  
FT /note= "minimal zinc binding domain"  
FT Domain 741..765  
FT /note= "transmembrane domain"  
FT Domain 766..805  
FT /note= "cytoplasmic domain"  
XX WO200018899-A2.  
PN 06-APR-2000.  
XX  
XX  
XX 29-SEP-1999; 99WO-US22976.  
PF  
XX 30-SEP-1998; 98US-0163648.  
PR  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Acton LS, Robison KE, Hsieh FY;  
PI  
XX WPI; 2000-293140/25.  
DR N-PSDB; AAA12764.  
XX  
XX Isolated nucleic acid encoding angiotensin converting enzyme-2 (ACE-2)  
PT polypeptide useful for detecting an ACE-2 therapeutic for treating  
PT hypertension, congestive heart failure, myocardial infarction,  
PT atherosclerosis and renal failure -  
XX  
XX Claim 2; Fig 1; 138pp; English.  
XX  
XX The present sequence represents a human angiotensin converting enzyme-2  
CC (ACE-2). ACE-2 is expressed predominantly in kidneys and testis. The  
CC sequence of the full length ACE-2 cDNA was determined from a clone  
CC obtained from a cDNA library prepared from mRNA of a human heart of  
CC a subject who had congestive heart failure. ACE-2 has significant  
CC sequence homologies with ACE enzymes, and has also been shown to  
CC hydrolyse angiotensin I into Ang.(1-9). The ACE-2 therapeutics are  
CC used to treat blood pressure related diseases and conditions, such as  
CC hypertension, congestive heart failure, chronic heart failure, acute  
CC heart failure, myocardial infarction, atherosclerosis and renal  
CC failure.  
XX  
XX Sequence 805 AA;  
SQ  
  
Query Match 5.0%; Score 8; DB 21; Length 805;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 81 LFGEEDVR 88  
Db 664 lfgeedvr 671  
  
RESULT 29  
AAAY67310  
ID AAY67310 standard; Protein; 805 AA.  
XX  
XX AAY67310;  
XX  
XX 11-APR-2000 (first entry)  
DT  
XX Human MPROT15 amino acid sequence #1.  
DE  
XX MPROT15; treatment; hypertension; human; myocardial disease; apoplexy;  
KW heart disease; apoplexy; heart disease; nervous denaturation; hormone;  
KW Alzheimer's disease; cytokine.  
XX Homo sapiens.  
OS

XX JP11318472-A.  
PN  
XX 24-NOV-1999.  
PD  
XX 22-JAN-1999; 99JP-0014949.  
PF  
XX 13-MAY-1998; 98GB-0010373.  
PR 18-AUG-1998; 98GB-0018009.  
XX  
XX (SMIX ) SMITHKLINE BEECHAM PLC.  
PA  
XX WPI; 2000-109268/10.  
XX N-PSDB; AA259465.  
DR  
XX MPROT15 polypeptide and MPROT15 polynucleotides - useful for the  
PT treatment of hypertension, myocardial diseases, apoplexy, heart  
PT diseases, nervous denaturation, Alzheimer's disease etc.  
XX  
XX Claim 1; Page 15; 22pp; Japanese.  
PS  
XX This is amino acid sequence #1 of human MPROT15. The MPROT15  
CC polynucleotide and polypeptide sequences can be used for the treatment of  
CC hypertension, myocardial diseases, apoplexy, heart diseases, nervous  
CC denaturation, Alzheimer's disease and diseases related to the processing  
CC of peptide hormones and cytokines.  
XX  
XX Sequence 805 AA;  
SQ  
  
Query Match 5.0%; Score 8; DB 21; Length 805;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 81 LFGEEDVR 88  
Db 664 lfgeedvr 671  
  
RESULT 30  
AAAY72667  
ID AAY72667 standard; Protein; 805 AA.  
XX  
XX AAY72667;  
XX  
XX 31-MAY-2001 (first entry)  
DT  
XX Human angiotensin converting enzyme-2 (ACE-2).  
DE  
XX  
XX Human; angiotensin converting enzyme-2; ACE-2; peptidyl dipeptidase A;  
KW screening; therapy; hypertension; congestive heart failure; CHF;  
KW inflammation; pain.  
XX  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH Peptide 1..18  
FT /label= Signal\_peptide  
FT Protein 15..805  
FT /label= Mature\_ACE-2\_protein  
FT Domain 374..378  
FT /label= ZBD  
FT /note= "Zinc binding domain"  
FT Domain 741..765  
FT /label= TMD  
FT /note= "Transmembrane domain; Hydrophobic region"  
FT Domain 766..805  
FT /label= Cytoplasmic\_domain  
XX  
XX US6194556-B1.  
PN  
XX 27-FEB-2001.  
PD  
XX

PF 11-DEC-1997; 97US-0989299.  
 XX  
 PR 11-DEC-1997; 97US-0989299.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Acton SL, Robison KE;  
 PI  
 XX  
 XX WPI: 2001-210604/21.  
 DR N-PSDB; AAD02758.  
 DR

XX Novel genes encoding angiotensin converting enzyme-2 useful as  
 PT antisense or antigenic agents for therapeutics, diagnostics and  
 PT screening assays  
 PT

PS Claim 33; Fig 1; 76pp; English.

CC The present amino acid sequence is human angiotensin converting enzyme-2  
 CC (ACE-2), also referred as peptidyl dipeptidase A (EC 3.4.15.1). Nucleic  
 CC acid sequence encoding ACE-2 is useful as antisense or antigenic agents  
 CC for sequence specific modulation of gene expression or in the analysis of  
 CC single base-pair mutations in the gene. Nucleic acid sequence encoding  
 CC ACE-2 is useful in therapeutics, diagnostics and in screening assays.  
 CC ACE-2 antagonist is used to treat hypertension or congestive heart  
 CC failure (CHF). ACE agonist is used to reduce the inflammation and pain  
 CC resulting from an insect sting or bite, which was accompanied by an  
 CC injection of bradykinin. Anti-ACE-2 antibodies are used to monitor ACE-2  
 CC protein levels for determining the disease or condition associated with  
 CC an aberrant protein level.  
 XX

SQ Sequence 805 AA;

Query Match 5.0%; Score 8; DB 22; Length 805;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LFGEEDVR 88  
 |||||  
 Db 664 lfgeedvr 671

RESULT 31

AAB48095  
 ID AAB48095 standard; Protein; 805 AA.

AC AAB48095;

XX 19-MAR-2001 (first entry)

DE Human Zace2 protein.

XX Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;  
 KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;  
 KW ventricular systolic dysfunction; renal impairment; heart failure;  
 KW scleroderma renal crisis; atherosclerosis; antiinflammatory; human;  
 KW antiarthritic; bradykinin inactivator.  
 XX

OS Homo sapiens.

XX WO200070032-A1.

PN 23-NOV-2000.

XX 03-MAY-2000; 2000WO-US11932.

PF 13-MAY-1999; 99US-0311482.

PR 27-AUG-1999; 99US-0384706.

XX (ZYMO) ZYMOGENETICS INC.

XX Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;

XX

DR WPI: 2001-025018/03.  
 DR N-PSDB; AAC84366, AAC84367.

XX Angiotensin-converting enzyme, Zace2, useful for treating inflammatory  
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases  
 PT associated with inflammation such as arthritis and enterocolitis -  
 XX  
 XX Example 1; Page 95-100; 125pp; English.

XX The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-  
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood  
 CC pressure regulation and fertility. Zace2 can be expressed by standard  
 CC recombinant methodology. Zace2 polypeptides are useful for treating an  
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),  
 CC diseases associated with inflammation like arthritis and enterocolitis,  
 CC as targets for identifying modulators of zinc protease activity, for  
 CC screening or identifying new angiotensin-converting enzyme (ACE)  
 CC inhibitors, and as a basis for rational drug design for inhibitory  
 CC molecules. The nucleic acids can be used to detect the expression of a  
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and  
 CC for detecting and localizing Zace2 gene expression in tissue samples,  
 CC to determine whether a subject's chromosomes contain a mutation in the  
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.  
 CC Inhibitors of ACE are used for treating hypertension of various  
 CC conditions, including left ventricular systolic dysfunction, progressive  
 CC renal impairment, scleroderma renal crisis, congestive heart failure due  
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be  
 CC used to treat infertility while Zace2 antagonists are used for inducing  
 CC infertility. The present sequence represents the human Zace2 protein.  
 XX

SQ Sequence 805 AA;

Query Match 5.0%; Score 8; DB 22; Length 805;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LFGEEDVR 88  
 |||||  
 Db 664 lfgeedvr 671

RESULT 32

AAW17971  
 ID AAW17971 standard; Protein; 47 AA.

AC AAW17971;

XX 12-DEC-1997 (first entry)

XX RAC protein kinase C-terminal binding protein C-terminal region.

XX RAC protein kinase C-terminal binding protein; CTBP; modulator;  
 KW signal transduction; insulin; cell proliferation; glycogen.

OS Homo sapiens.

XX WO9718303-A1.

XX 22-MAY-1997.

XX 05-NOV-1996; 96WO-EP04810.

PR 15-DEC-1995; 95GB-0025704.

PR 16-NOV-1995; 95GB-0023379.

XX (NOVS) NOVARTIS AG.

XX Hemmings BA;

XX WPI: 1997-289279/26.

DR N-PSDB; AAT67134.

XX

PT RAC protein kinase, or modulator excluding wortmannin and vanadate,  
 PT for use as medicament - and screening potential modulators of  
 PT insulin mediated intracellular signalling using RAC-PK, or fragment  
 XX  
 PS  
 PS Claim 8; Page 25; 38pp; English.  
 XX  
 CC This polypeptide comprises the putative C-terminal region of a  
 CC novel RAC protein kinase C-terminal binding protein (CTBP). Its  
 CC sequence was deduced from a cDNA clone (AA167134) isolated from a  
 CC HeLa library following a yeast two-hybrid screen assay for clones  
 CC that showed specific interaction with RAC's kinase domain with its  
 CC C-terminal extension. The C-terminal domain of RAC protein kinase  
 CC is phosphorylated in response to insulin activation, suggesting a  
 CC role for CTBP as a modulator of insulin action. RAC protein kinase  
 CC signalling such as CTBP can be used in the treatment of  
 CC abnormalities of cellular metabolism, diseases involving an anomaly  
 CC in insulin response, and diseases involving an anomaly in glycogen  
 CC metabolism.  
 XX  
 SQ Sequence 47 AA;  
 SQ  
 Query Match 4.4%; Score 7; DB 18; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NSARARA 7  
 Db | | | | | | |  
 1 nsarara 7  
 RESULT 33  
 AAB32429  
 ID AAB32429 standard; Protein; 51 AA.  
 XX  
 AC AAB32429;  
 XX  
 DT 16-JAN-2001 (first entry)  
 XX  
 DE Human secreted protein sequence encoded by gene 6 SEQ ID NO:115.  
 XX  
 KW Human; secreted protein; cytostatic; immunostimulant; antiproliferative;  
 KW cardiant; antiarrhythmic; antiviral; antibacterial; antifungal; cancer;  
 KW antiparasitic; neuroprotective; nootropic; antiinflammatory; anti-HIV;  
 KW antiangiogenic; antiarteriosclerotic; diagnosis; immune disorder; AIDS;  
 KW autoimmune disease; haematopoietic cell disorder; blood protein disorder;  
 KW agammaglobulinaemia; hyperproliferative disease; Gaucher's disease;  
 KW cardiovascular disorder; congenital heart defect; pulmonary atresia;  
 KW arrhythmia; ischaemia; angiogenesis related disorder; Crohn's disease;  
 KW atherosclerosis; neurological disease; Alzheimer's disease;  
 KW Huntington's; infectious disease; cat-scratch disease.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200047602-A1.  
 PN  
 XX 17-AUG-2000.  
 XX  
 XX 08-FEB-2000; 2000WO-US03062.  
 PF  
 XX 10-FEB-1999; 99US-0119468..  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Soppet DR, Moore PA;  
 PI Shi Y, Lafleur DW, Olsen HS, Florence KA, Komatsoulis G;  
 DI  
 XX WPI: 2000-543578/49.  
 DR  
 XX New human nucleic acids encoding secreted proteins, useful in the  
 XX treatment, prevention or diagnosis of immune disorders (e.g. autoimmune  
 PT diseases), blood protein disorders and hyperproliferative diseases  
 PT

PT (e.g. Gaucher's disease) -  
 XX Disclosure; Page 18; 488pp; English.  
 XX  
 CC The polynucleotide sequences given in AAC55190 to AAC55235 encode the  
 CC human secreted proteins given in AAB32371 to AAB32484. Human secreted  
 CC proteins have activities based on the tissues and cells the genes are  
 CC expressed in. Examples of activities include: cytostatic;  
 CC immunostimulant; antiproliferative; cardiant; antiarrhythmic; antiviral;  
 CC antibacterial; antifungal; antiparasitic; neuroprotective; nootropic;  
 CC antiinflammatory; angiogenic; anti-HIV; and antiarteriosclerotic. The  
 CC polynucleotides and polypeptides, or their agonists and antagonists, can  
 CC be used for treating, preventing or diagnosing immune disorders (e.g.  
 CC cancer, autoimmune diseases), disorders of haematopoietic cells, blood  
 CC protein disorders (e.g. agammaglobulinaemia), hyperproliferative  
 CC diseases (e.g. Gaucher's disease), cardiovascular disorders (e.g.  
 CC congenital heart defects, pulmonary atresia, arrhythmias, ischaemia),  
 CC angiogenesis related disorders (e.g. Crohn's disease, atherosclerosis),  
 CC neurological diseases (e.g. Alzheimer's disease, Huntington's chorea),  
 CC infectious diseases (e.g. AIDS, cat-scratch disease and other bacterial,  
 CC viral, parasitic or fungal diseases). AAC55181 to AAC55189 and AAB32370  
 CC represent sequences used in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 51 AA;  
 SQ  
 Query Match 4.4%; Score 7; DB 21; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NSARARA 7  
 Db | | | | | | |  
 4 nsarara 10  
 RESULT 34  
 AAB02095  
 ID AAB02095 standard; Protein; 72 AA.  
 XX  
 AC AAB02095;  
 XX  
 DT 03-JAN-2001 (first entry)  
 XX  
 DE Cyclohexanone monooxygenase sequence.  
 DE  
 DE Identification; prokaryote; polymerase chain reaction; PCR;  
 KW amplification; primer; differential display;  
 KW picric acid degradation; gene cluster; open reading frame; ORF;  
 KW dehydratase; dehydrogenase; transcription factor;  
 KW Acyl-CoA synthase; NADPH oxidoreductase;  
 KW cyclohexanone monooxygenase.  
 XX  
 OS Unknown.  
 XX  
 XX WO200049177-A2.  
 PN  
 XX 24-AUG-2000.  
 XX  
 XX 17-FEB-2000; 2000WO-US03989.  
 PF  
 XX 19-FEB-1999; 99US-0120702.  
 PR  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PA  
 XX Rouviere P;  
 PI  
 XX WPI: 2000-587069/55.  
 DR  
 XX Differential display method using a large number of arbitrary primers  
 PT for RT-PCR used to isolate novel differentially expressed prokaryotic  
 PT genes  
 PT  
 XX

PS Example 8; Page 65; 66pp; English.

XX A new method to identify differentially expressed prokaryotic genes  
CC using a large number of arbitrarily primed polymerase chain  
CC reactions comprising separating two populations of microbial cells,  
CC where a first population is contacted with a stimulating agent;  
CC extracting total RNA from both microbial cell populations;  
CC amplifying the extracted RNA from both populations by preparing a  
CC collection of at least thirty-two different arbitrary primers, where  
CC each primer comprises a common and a variable region; individually  
CC contacting each primer of with a sample of extracted RNA from the  
CC two populations under conditions where two sets of amplification  
CC products are produced; purifying the two sets of amplification  
CC products; identifying the amplification products generated in the  
CC first population which differ from products generated from the  
CC second population as differentially expressed genes; and optionally  
CC sequencing the identified differentially expressed genes. The  
CC advantage over previous methods is that previous methods of  
CC differential display to clone genes using thirty-two or thirty  
CC primers have isolated four and one genes, respectively. The  
CC new method using a greater number of primers has isolated  
CC twenty-one induced gene fragments.

XX Sequence 72 AA;

Query Match 4.4%; Score 7; DB 21; Length 72;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 GAKLLG 152  
| | | | |  
DB 17 gaklllg 23

RESULT 35

AAY14475

ID AAY14475 standard; Protein; 74 AA.

XX AAY14475;

XX 17-AUG-1999 (first entry)

DE Fragment of human secreted protein encoded by gene 7.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

XX WO9919339-A1.

XX 22-APR-1999.

XX 08-OCT-1998; 98WO-US21142.

XX 09-OCT-1997; 97US-0071498.

PR 09-OCT-1997; 97US-0061463.

PR 09-OCT-1997; 97US-0061527.

PR 09-OCT-1997; 97US-0061529.

PR 09-OCT-1997; 97US-0061532.

PR 09-OCT-1997; 97US-0061536.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Brewer LA, Duan R, Ebner R, Ferrie AM, Florence C;

PI Florence KA, Greene JM, Olsen HS, Rosen CA, Ruben SM;

PI Young PE, Yu G;

XX WPI; 1999-277587/23.

XX New isolated human genes and the secreted polypeptides they encode

XX Disclosure; Page 14; 226pp; English.

XX This sequence represents a fragment of a secreted human protein encoded  
CC by the nucleic acid molecule detailed in the descriptor line. The gene  
CC can be used to generate fusion proteins by linking to the gene to a  
CC human immunoglobulin Fc portion (e.g. AAX79002) for increasing the  
CC stability of the fused protein as compared to the human protein only.  
CC The invention relates to 53 novel genes and their fragments (nucleic  
CC acid sequences: AAX79011-X79064; amino acid sequences AAY14411-Y14464)  
CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC conditions can be diagnosed by determining the amount of the new  
CC polypeptides in a sample or by determining the presence of mutations in  
CC the new polynucleotides. Specific uses are described for each of the 53  
CC polynucleotides, based on which tissues they are most highly expressed in  
CC (see AAX79011 for described uses).

XX Sequence 74 AA;

Query Match 4.4%; Score 7; DB 20; Length 74;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSARARA 7  
| | | | |  
DB 39 nsarara 45

RESULT 36

AAB53775

ID AAB53775 standard; Protein; 84 AA.

XX AAB53775;

XX 09-MAR-2001 (first entry)

DE Human colon cancer antigen protein sequence SEQ ID NO:1315.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW identification; cytostatic; cardioactive; neuroprotective; vulnary;  
KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;  
KW neural disorder; immune system disorder; muscular disorder;  
KW reproductive disorder; gastrointestinal disorder; renal disorder;  
KW infectious disease; cardiovascular disorder.

XX Homo sapiens.

XX WO200055351-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05883.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587534/55.

XX N-PSDB; AAC98532.

XX Colon cancer associated gene sequences, referred to as colon cancer  
PT antigens, useful for the treatment, prevention, and diagnosis of colon  
PT disorders such as colon cancer -

XX Claim 11; Page 1897; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,

CC called human colon cancer antigens, given in AAB53234 to AAB54006. The

CC human colon cancer antigens can have cytostatic, cardioactive, muscular;

CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,

CC vulnary, nephrotropic, antiinfective and antibacterial activities, and

CC can be used in gene therapy. The colon cancer antigen polynucleotides,

CC proteins and antibodies to the proteins are useful for the prevention,

CC treatment and diagnosis of colon disorders, such as colon cancer. The

CC polynucleotides may be used in diagnostics and research, such as for

CC chromosome identification, and as hybridisation probes. The proteins

CC may also be used to prevent diseases such as neural disorders, immune

CC system disorders, muscular disorders, reproductive disorders,

CC gastrointestinal disorders, wounds, renal disorders, infectious

CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and

CC AAB54007 represent sequences used in the exemplification of the present

XX invention.

XX Sequence 84 AA;

Query Match 4.4%; Score 7; DB 21; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ARARAVL 9  
 |||||

Db 53 araravl 59

RESULT 37  
 AAY53890

ID AAY53890 standard; Protein; 87 AA.

XX AC AAY53890;

DT 13-MAR-2000 (first entry)

XX Partial amino acid sequence of human interleukin-21.

XX Human; interleukin-21; IL-21; IL-22; immune system disorder;  
 KW immune cell chemotaxis; haematopoietic cell disorder;  
 KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;  
 KW allergic asthma; respiratory problem; organ rejection;  
 KW graft-versus-host disease; GVHD; inflammation;  
 KW hyperproliferative disorder; tissue regeneration;  
 KW embryonic stem cell differentiation; embryonic stem cell proliferation;  
 KW haematopoietic lineage.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Domain 3..11 /note= "conserved domain I"

FT Domain 19..24 /note= "conserved domain II"

FT Domain 46..52 /note= "conserved domain III"

FT Domain 75..82 /note= "conserved domain IV"

XX WO9961617-A1.

XX 02-DEC-1999.

XX 27-MAY-1999; 99WO-US11644.

XX 29-MAY-1998; 98US-0087340.

PR 10-SEP-1998; 98US-0099805.

PR 30-APR-1999; 99US-0131965.

XX

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Ebner R;

WPI; 2000-072622/06.

N-PSDB; AAZ36834.

Novel polynucleotides used to develop products for treating e.g. immune disorders, blood disorders, autoimmune disorders, allergies, inflammation, hyperproliferative disorders or infections -

Claim 25; Fig 1; 170pp; English.

The present sequence represents a partial human interleukin-21 (IL-21) protein. The specification also describes IL-22 polynucleotides and polypeptides. The IL-21 polynucleotide was isolated from a cDNA library of apoptotic T-cells. IL-21 and IL-22 may be useful in treating or inhibiting the proliferation of the immune system, by activating or (chemotaxis) of immune cells, treating or detecting deficiencies or disorders of haematopoietic cells, to modulate haemostatic or thrombolytic activity, in treating or detecting autoimmune disorders, treating asthma (particularly allergic asthma) or other respiratory problems, to treat and/or prevent organ rejection or graft-versus-host disease (GVHD), to modulate inflammation, to treat or detect hyperproliferative disorders, to treat or detect infectious agents, to differentiate, proliferate and attract cells, leading to the regeneration of tissues. IL-21 and IL-22 may also increase or decrease the differentiation or proliferation of embryonic stem cells and haematopoietic lineage, may be used to modulate mammalian characteristics.

Sequence 87 AA;

Query Match 4.4%; Score 7; DB 21; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PVGCTCV 128  
 |||||

Db 76 pvgctcv 82

RESULT 38  
 AAB56590

ID AAB56590 standard; Protein; 90 AA.

XX AC AAB56590;

XX 13-MAR-2001 (first entry)

XX Human prostate cancer antigen protein sequence SEQ ID NO:1168.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis; neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular; vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological; antibacterial; gene therapy; immune; reproductive; renal; gastrointestinal; pulmonary; cardiovascular; proliferative disorder; wound; infectious disease.

XX Homo sapiens.

XX WO200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05988.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

```

XX PI Rosen CA, Ruben SM;
XX
XX WPI; 2000-587513/55.
DR N-PSDB; AAF15793.
XX
XX Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
XX Claim 11; Page 1575; 2338pp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
XX Sequence 90 AA;
SQ

```

Query Match 4.4%; Score 7; DB 21; Length 90;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 RARAVLS 10
Db 8 raravls 14

```

RESULT 39  
AAB52552  
ID AAB52552 standard; Protein; 114 AA.  
XX  
XX AAB52552;  
XX  
XX 23-FEB-2001 (first entry)  
XX  
XX Helicobacter pylori bait polypeptide #70.  
XX  
XX Helicobacter pylori; two-hybrid system; protein-protein interaction;  
KW bait polypeptide; gastric ulcer; antibacterial.  
XX  
XX Helicobacter pylori.  
OS  
XX WO200066722-A1.  
XX  
XX 09-NOV-2000.  
XX  
XX 14-APR-2000; 2000WO-IB00603.  
XX  
XX 30-APR-1999; 99EP-0401066.  
XX  
XX (HYBR-) HYBRIGENICS SA.  
XX  
XX Legrain P, Selig L, Rain J;  
XX  
XX WPI; 2000-687535/67.  
XX N-PSDB; AAC97298.  
XX  
XX A two-hybrid system for identifying compounds useful in the treatment  
PT of e.g. gastric ulcers comprises producing a collection of recombinant  
XX cell clones -  
XX

```

PS Example 5; Page 167-168; 267pp; English.
XX
XX The present sequence is a bait polypeptide used in a Helicobacter
CC pylori two-hybrid screen to identify protein-protein interactions.
CC The method is used to identify a recombinant cell clone expressing a
CC prey polypeptide which is capable of interacting with the bait
CC polypeptide. The two hybrid system is useful for screening compounds
CC for antibacterial activity. It may be used in the treatment of gastric
CC ulcers. The polynucleotides are useful as amplification primers or
CC specific detection probes. The polypeptides, vectors or host cells can
CC be used as immunogens to produce mono- or polyclonal antibodies. The
CC polynucleotides, polypeptides, antibodies, vectors, host cells or
CC modulating agents can be used to produce a pharmaceutical composition.
XX
XX Sequence 114 AA;
SQ

```

Query Match 4.4%; Score 7; DB 21; Length 114;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 80 GLRGEED 86
Db 40 glrgeed 46

```

RESULT 40  
AAB07601  
ID AAB07601 standard; Protein; 123 AA.  
XX  
XX AAB07601;  
XX  
XX 07-NOV-2000 (first entry)  
XX  
XX A human interleukin (IL) 171 polypeptide.  
DE Interleukin; IL-17; CTLA-8; IL-170; IL-172; IL-173; IL-174; IL-176;  
XX IL-177; IL-171; cell proliferation; cancer.  
XX Homo sapiens.  
XX WO2000042188-A2.  
XX 20-JUL-2000.  
XX 10-JAN-2000; 2000WO-US000006.  
XX  
XX 11-JAN-1999; 99US-0228822.  
XX  
XX (SCHE ) SCHERING CORP.  
XX  
XX Gorman DM, Bazan JF, Kastelein RA;  
XX  
XX WPI; 2000-466130/40.  
XX N-PSDB; AAA58990.  
XX  
XX New isolated polynucleotide encoding a mammalian Interleukin-17 like  
PT protein used to identify genes for homologous proteins -  
XX  
XX Disclosure; Page 19-20; 111pp; English.  
XX  
XX The present sequence represents an interleukin-171 (IL-171) polypeptide.  
CC The polypeptide is an IL-17-like (CTLA-8 related) protein. It is a  
CC member of a new group of interleukins, IL-170 polypeptides. The members  
CC comprise IL-172, IL-173, IL-174, IL-176, IL-177, and IL-171. IL-170  
CC protein can be used to treat abnormal proliferation e.g. cancer  
CC or degenerative conditions. Antibodies can be used in diagnostic  
CC methods to detect over production of IL-170 protein in cells or body  
CC fluids.  
XX  
XX Sequence 123 AA;  
SQ

Query Match 4.4%; Score 7; DB 21; Length 123;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PVGCTCV 128  
Db 73 pvgctcv 79

RESULT 41  
AAB07683  
ID AAB07683 standard; Protein; 123 AA.  
XX  
AC AAB07683;  
XX  
DT 07-NOV-2000 (first entry)  
XX  
DE A human interleukin-171 polypeptide.  
XX  
KW Interleukin; IL-171; cytokine; CTLA-8; IL-17; IL-175; IL-172; IL-173;  
KW IL-174; IL-176; IL-177; cell proliferation; cancer.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 94  
FT Misc-difference 123 /note= "unspecified amino acid encoded by GNT"  
FT Misc-difference 123 /note= "unspecified amino acid encoded by NGC"  
XX  
PN WO200042187-A1.  
XX  
PD 20-JUL-2000.  
XX  
PF 10-JAN-2000; 2000WO-US000005.  
XX  
PR 11-JAN-1999; 99US-0229402.  
XX  
PA (SCHE ) SCHERING CORP.  
XX  
PI Gorman DW, Bazan JF, Kastelein RA;  
XX  
DR WPI; 2000-476060/A1.  
DR N-PSDB; AAA59148.  
XX  
XX  
XX New DNA sequence encoding a mammalian homolog of CTLA-8, designated  
XX Interleukin-171 (IL-171), useful for recombinant production of IL-171  
XX which can be used for treating conditions associated with abnormal  
XX physiology or development -  
XX  
PS Claim 11; Page 9-10; 111pp; English.  
XX  
XX The present sequence represents an interleukin (IL)-171 polypeptide.  
XX It is a mammalian homologue of the cytokine designated CTLA-8 (also  
XX referred to as IL-17). The specification also describes homologues  
XX IL-171, IL-175, IL-172, IL-173, IL-174, IL-176, and IL-177. The DNA  
XX sequence encoding IL-171 is useful for identifying genes, mRNA and  
XX cDNA molecules which code for related or homologous proteins. The  
XX IL-171 protein, antibodies against IL-171, and compounds which have  
XX binding affinity to IL-171 are useful in treatment of conditions  
XX associated with abnormal physiology or development, including abnormal  
XX proliferation, e.g. cancerous conditions, or degenerative conditions.  
XX The IL-171 protein can be used in kits and assay methods for identifying  
XX compounds that selectively bind to IL-171.  
XX  
SQ Sequence 123 AA;

Query Match 4.4%; Score 7; DB 21; Length 123;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PVGCTCV 128

Db 73 pvgctcv 79

RESULT 42  
AAW98374  
ID AAW98374 standard; Protein; 170 AA.  
XX  
AC AAW98374;  
XX  
DT 31-MAR-1999 (first entry)  
XX  
DE H. pylori GHPO 1120 protein.  
XX  
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
KW peptic ulcer disease.  
XX  
OS Helicobacter pylori.  
XX  
PN WO9843478-A1.  
XX  
PD 08-OCT-1998.  
XX  
PF 01-APR-1998; 98WO-US06371.  
XX  
PR 29-JUL-1997; 97US-0902615.  
PR 01-APR-1997; 97US-0833457.  
PR 24-JUN-1997; 97US-0881227.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
XX  
PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
XX  
DR WPI; 1998-542293/46.  
DR N-PSDB; AAX14093.  
XX  
XX New isolated Helicobacter polynucleotides - used to develop products  
XX for the diagnosis, prevention and treatment of Helicobacter  
XX infections and gastrointestinal diseases  
XX  
PS Claim 8; Page 569-570; 2054pp; English.  
XX  
XX This sequence represents a Helicobacter pylori GHPO protein of the  
XX invention. The polypeptides can be used for preventing or treating  
XX Helicobacter infections, and gastroduodenal diseases associated with  
XX these infections, including acute, chronic, and atrophic gastritis, and  
XX peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
XX used for the production of antibodies. The products can also be used for  
XX detection and diagnosis.  
XX  
SQ Sequence 170 AA;

Query Match 4.4%; Score 7; DB 19; Length 170;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 EKDADSI 138  
Db 69 ekdads1 75

RESULT 43  
AAB18911  
ID AAB18911 standard; Protein; 197 AA.  
XX  
AC AAB18911;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE A novel polypeptide designated PRO1122.

KW Secreted protein; transmembrane protein; PRO1484; PRO4334; PRO1122;  
KW PRO1889; PRO1887; PRO1785; PRO4353; PRO4405; PRO4356;  
KW PRO4352; PRO4380; PRO4354; PRO4408; PRO5737; PRO4425; PRO5990; PRO6030;  
KW PRO4424; PRO4422; PRO4430; PRO4499; tumour; obesity; diabetes;  
KW insulinemia; kidney disorder; Bergers disease; nephropathy;  
KW Schonlein-Henoch purpura; celiac disease; dermatitis herpetiformis;  
KW Crohns disease.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Region 3..25  
FT /note= "signal peptide"  
FT /note= "leucine zipper pattern"  
FT Modified-site 32..38  
FT /note= "N-myristoylation site"  
FT Modified-site 55..61  
FT /note= "N-myristoylation site"  
FT Modified-site 112..121  
FT /note= "tyrosine kinase phosphorylation site"  
FT Modified-site 133..139  
XX  
XX WO200056889-A2.  
XX  
XX 28-SEP-2000.  
XX  
XX 01-MAR-2000; 2000WO-US05601.  
XX  
XX 23-MAR-1999; 99US-0125774.  
XX 23-MAR-1999; 99US-0125778.  
XX 24-MAR-1999; 99US-0125826.  
XX 31-MAR-1999; 99US-0127035.  
XX 05-APR-1999; 99US-0127706.  
XX 21-APR-1999; 99US-0130359.  
XX 27-APR-1999; 99US-0131270.  
XX 27-APR-1999; 99US-0131272.  
XX 27-APR-1999; 99US-0131291.  
XX 04-MAY-1999; 99US-0132371.  
XX 04-MAY-1999; 99US-0132379.  
XX 04-MAY-1999; 99US-0132383.  
XX 25-MAY-1999; 99US-0135750.  
XX 08-JUN-1999; 99US-0138166.  
XX 20-JUL-1999; 99US-0144791.  
XX 03-AUG-1999; 99US-0146970.  
XX 09-DEC-1999; 99US-0170262.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Desnoyers L, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J;  
PI Stewart TA, Watanabe CK, Wood WI, Zhang Z;  
XX  
XX WPI: 2000-628263/60.  
XX N-PSDB; AAA96338.  
XX  
XX Novel secreted and transmembrane polypeptides useful for diagnosing  
PT tumour in a mammal, for identifying agonists and antagonists of the  
PT polypeptide and for therapeutic use  
XX  
XX Claim 12; Fig 6; 222pp; English.  
XX  
XX The present sequence represents a secreted or transmembrane polypeptide.  
CC The specification describes polypeptides designated PRO1484, PRO4334,  
CC PRO1122, PRO1889, PRO1890, PRO1785, PRO4353, PRO4357, PRO4405,  
CC PRO4356, PRO4352, PRO4380, PRO4354, PRO4408, PRO5737, PRO4425, PRO5990,  
CC PRO6030, PRO4424, PRO4422, PRO4430 and PRO4499. PRO1889 polypeptide is  
CC useful for diagnosing tumour in a mammal. The polypeptides, their  
CC agonists and antagonists are useful treating a condition associated with  
CC expression or activity of the polypeptide. Conditions treated include  
CC obesity, diabetes or hyper- or hypo-insulinemia. The polypeptides are  
CC capable of inducing proliferation of mammalian kidney mesangial cells  
CC and are therefore useful for treating kidney disorders associated with  
CC decreased mesangial cell function such as Bergers disease or other

CC nephropathies associated with Schonlein-Henoch purpura, celiac disease,  
CC dermatitis herpetiformis or Crohns disease. The nucleic acids may be used  
CC to generate transgenic animals for use in development and screening of  
CC therapeutically useful reagents and also for chromosome identification  
CC and tissue typing.  
XX  
XX Sequence 197 AA;  
SQ  
  
Query Match 4.4%; Score 7; DB 21; Length 197;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 122 PVGCTCV 128  
Db 186 pvgctcv 192  
|||||||  
  
RESULT 44  
AAB07602  
ID AAB07602 standard; Protein; 197 AA.  
XX  
XX AAB07602;  
AC  
XX  
XX 07-NOV-2000 (first entry)  
XX  
XX A human interleukin (IL) 171 polypeptide.  
DE  
XX  
XX Interleukin; IL-17; CTLA-8; IL-170; IL-172; IL-173; IL-174; IL-176;  
KW IL-177; IL-171; cell proliferation; cancer.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FT Peptide 1..17  
FT /note= "signal peptide"  
FT Protein 18..197  
FT /note= "mature protein"  
XX  
XX WO200042188-A2.  
XX  
XX 20-JUL-2000.  
XX  
XX 10-JAN-2000; 2000WO-US000006.  
XX  
XX 11-JAN-1999; 99US-0228822.  
XX  
XX (SCHE ) SCHERING CORP.  
XX  
XX Gorman DM, Bazan JF, Kastelein RA;  
XX  
XX WPI: 2000-466130/40.  
XX N-PSDB; AAA58991.  
XX  
XX New isolated polynucleotide encoding a mammalian Interleukin-17 like  
PT protein used to identify genes for homologous proteins -  
XX  
XX Disclosure; Page 20-21; 111pp; English.  
XX  
XX The present sequence represents an interleukin-171 (IL-171) polypeptide.  
CC The polypeptide is an IL-17-like (CTLA-8 related) protein. It is a  
CC member of a new group of interleukins, IL-170 polypeptides. The members  
CC comprise IL-172, IL-173, IL-174, IL-176, and IL-171. IL-170  
CC protein can be used to treat abnormal proliferation e.g. cancer  
CC or degenerative conditions. Antibodies can be used in diagnostic  
CC methods to detect over production of IL-170 protein in cells or body  
XX fluids.  
XX  
XX Sequence 197 AA;  
SQ  
  
Query Match 4.4%; Score 7; DB 21; Length 197;  
Best Local Similarity 100.0%; Pred. No. 37;

```
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PVGCTCV 128
    |||||
Db 186 pvgctcv 192

RESULT 45
AAB07684
ID AAB07684 standard; Protein; 197 AA.
AC
XX
XX AAB07684;
DT
XX
XX 07-NOV-2000 (first entry)
DE
XX
XX A human interleukin-171 polypeptide.
KW Interleukin; IL-171; cytokine; CTLA-8; IL-17; IL-175; IL-172; IL-173;
KW IL-174; IL-176; IL-177; cell proliferation; cancer.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Peptide 1..17
XX FT /note= "signal peptide"
XX FT Protein 18..197
XX FT /note= "mature protein"
XX FT Modified-site 55..57
XX FT /note= "putative glycosylation site"
XX
XX WO200042187-A1.
XX
XX 20-JUL-2000.
XX
XX 10-JAN-2000; 2000WO-US000005.
XX
XX 11-JAN-1999; 99US-0229402.
XX
XX (SCHE ) SCHERING CORP.
XX
XX Gorman DM, Bazan JF, Kastelein RA;
XX
XX WPI; 2000-476060/A1.
XX N-PSDB; AAA59149.
XX
XX New DNA sequence encoding a mammalian homolog of CTLA-8, designated
XX interleukin-171 (IL-171), useful for recombinant production of IL-171
XX which can be used for treating conditions associated with abnormal
XX physiology or development -
XX
XX Claim 11; Page 10-11; 11lpp; English.
XX
XX The present sequence represents an interleukin (IL)-171 polypeptide.
XX It is a mammalian homologue of the cytokine designated CTLA-8 (also
XX referred to as IL-17). The specification also describes homologues
XX IL-171, IL-175, IL-173, IL-174, IL-176, and IL-177. The DNA
XX sequence encoding IL-171 is useful for identifying genes, mRNA and
XX cDNA molecules which code for related or homologous proteins. The
XX IL-171 protein, antibodies against IL-171, and compounds which have
XX binding affinity to IL-171 are useful in treatment of conditions
XX associated with abnormal physiology or development, including abnormal
XX proliferation, e.g. cancerous conditions, or degenerative conditions.
XX The IL-171 protein can be used in kits and assay methods for identifying
XX compounds that selectively bind to IL-171.
XX
XX Sequence 197 AA;

Query Match 4.4%; Score 7; DB 21; Length 197;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PVGCTCV 128
    |||||
Db 186 pvgctcv 192

RESULT 47
```

```
Db 186 pvgctcv 192
    |||||
RESULT 46
AAY92238
ID AAY92238 standard; Protein; 197 AA.
XX
XX AAY92238;
XX
XX 10-AUG-2000 (first entry)
DT
XX
XX Human interleukin-17 (IL-17) homologue.
DE
XX
XX Interleukin 17; IL-17; haematopoiesis; chemotherapy; cytostatic;
KW antianemic; cardiant; hemostatic; anti-inflammatory; anti-HIV.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Peptide 1..18
XX FT /label= signal_peptide
XX FT Protein 19..197
XX FT /label= mature_protein
XX
XX WO200020593-A1.
XX
XX 13-APR-2000.
XX
XX 30-SEP-1999; 99WO-US22678.
XX
XX 02-OCT-1998; 98US-0102883.
XX 01-DEC-1998; 98US-0110405.
XX 11-JUN-1999; 99US-0138910.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Glasebrook AL, Su EW, Wei J, Liu L;
XX
XX WPI; 2000-303778/26.
XX N-PSDB; AAA09153.
XX
XX Nucleic acid encoding an interleukin-17 (IL-17) homolog polypeptide
XX PT which enhances hematopoiesis, useful for treating e.g. anemia,
XX thrombocytopenia, viral and bacterial infections
XX
XX Claim 16; Page 92-93; 11lpp; English.
XX
XX Interleukin 17 (IL-17) stimulates hematopoiesis and production of
XX neutrophils, granulocytes, or platelets, this may be useful during
XX chemotherapy. IL-17 homologues have at least one activity selected
XX from induction of cytotoxic T cells, induction of lymphokine-activated
XX killer cell proliferation or a B or T cell stimulation. The IL-17
XX homologue may also be used to treat viral or bacterial infections,
XX immune related diseases, anemia, leukemia, thrombocytopenia, uremia,
XX Von Willebrand disease, postoperative cardiovascular dysfunction,
XX treatment of AIDS (acquired immune deficiency syndrome)-related bone
XX marrow failure, and inflammatory diseases of the gastrointestinal
XX system, joints, and lungs.
XX
XX Sequence 197 AA;

Query Match 4.4%; Score 7; DB 21; Length 197;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PVGCTCV 128
    |||||
Db 186 pvgctcv 192

RESULT 47
```

AAY44460  
 ID AAY44460 standard; Protein; 197 AA.  
 XX  
 AC AAY44460;  
 XX  
 DT 27-MAR-2000 (first entry)  
 XX  
 DE Human Interleukin 17C, PRO1122 polypeptide.  
 XX  
 KW Interleukin: IL-17C; PRO1122 polypeptide; clone DNA62377-1381-1; UNQ561;  
 KW cytokine IL-17; cytotoxic T-lymphocyte-associated antigen 8; CTLA-8;  
 KW hybridisation probe; antagonist; degenerative cartilaginous disorder;  
 KW agonist; diagnose; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..18  
 FT /label= Signal\_peptide  
 FT Protein 19..197  
 FT /label= Mature\_IL-17C\_polypeptide  
 FT /note= "used to treat degenerative cartilaginous  
 disorder"  
 FT Misc-difference 109  
 FT /note= "Conserved Trp residue"  
 FT Misc-difference 129  
 FT /note= "Conserved Cys residue"  
 FT Misc-difference 134  
 FT /note= "Conserved Cys residue"  
 FT Misc-difference 163  
 FT /note= "Conserved Cys residue"  
 FT Misc-difference 189  
 FT /note= "Conserved Cys residue"  
 FT Misc-difference 191  
 FT /note= "Conserved Cys residue"  
 FT  
 XX WO9960127-A2.  
 XX  
 XX 25-NOV-1999.  
 XX  
 XX 14-MAY-1999; 99WO-US10733.  
 XX  
 XX 15-MAY-1998; 98US-0085579.  
 XX 23-DEC-1998; 98US-0113621.  
 XX  
 XX (GETH ) GENENTECH INC.  
 XX  
 XX Chen J, Filvaroff E, Goddard A, Gurney AL, Li H, Wood WI;  
 XX WPI; 2000-116314/10.  
 XX N-PSDB; AAZ29728.  
 XX  
 XX New polypeptides designated PRO1031 and PRO1122 used to treat a  
 XX degenerative cartilaginous disorder -  
 XX  
 XX Claim 23; Fig 3; 141pp; English.  
 XX  
 XX The present sequence is the human PRO1122 polypeptide, also referred to  
 XX as UNQ561, and as interleukin-17C (IL-17C), encoded by the  
 XX clone DNA62377-1381-1. This sequence has identity with the  
 XX cytokine IL-17 and cytotoxic T-lymphocyte-associated antigen 8 (CTLA-8)  
 XX and has leucine zipper pattern. PRO1122 is expressed in pancreas, small  
 XX intestine, stomach and testis also. It shares about 26-28% amino acid  
 XX identity with IL-17 and IL-17B. The entire coding region of IL-17C can  
 XX be used as hybridisation probe. The PRO1122 polypeptide, agonist or  
 XX antagonist, is used to diagnose and treat a degenerative cartilaginous  
 XX disorder.  
 XX  
 XX Sequence 197 AA;  
 SQ  
 Query Match 4.4%; Score 7; DB 21; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 37;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 122 PVGCTCV 128  
 Db 186 pvgctcv 192  
 RESULT 48  
 AAY53892  
 ID AAY53892 standard; Protein; 197 AA.  
 XX  
 AC AAY53892;  
 XX  
 DT 13-MAR-2000 (first entry)  
 XX  
 DE Amino acid sequence of human interleukin-21.  
 XX  
 KW Human; Interleukin-22; IL-22; IL-21; immune system disorder;  
 KW immune cell chemotaxis; haematopoietic cell disorder;  
 KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;  
 KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;  
 KW inflammation; hyperproliferative disorder; tissue regeneration;  
 KW embryonic stem cell differentiation; embryonic stem cell proliferation;  
 KW haematopoietic lineage; allergic asthma.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..18  
 FT /note= "signal peptide"  
 FT Domain 34..40  
 FT /note= "conserved domain V"  
 FT Domain 63..68  
 FT /note= "conserved domain VI"  
 FT Domain 104..109  
 FT /note= "conserved domain VII"  
 FT Domain 113..121  
 FT /note= "conserved domain I"  
 FT Domain 129..134  
 FT /note= "conserved domain II"  
 FT Domain 156..162  
 FT /note= "conserved domain III"  
 FT Domain 185..192  
 FT /note= "conserved domain IV"  
 XX  
 XX WO9961617-A1.  
 XX  
 XX 02-DEC-1999.  
 XX  
 XX 27-MAY-1999; 99WO-US11644.  
 XX  
 XX 29-MAY-1998; 98US-0087340.  
 XX 10-SEP-1998; 98US-0099805.  
 XX 30-APR-1999; 99US-0131965.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Ebner R;  
 XX WPI; 2000-072622/06.  
 XX N-PSDB; AAZ36836.  
 XX  
 XX Novel polynucleotides used to develop products for treating e.g. immune  
 XX disorders, blood disorders, autoimmune disorders, allergies,  
 XX inflammation, hyperproliferative disorders or infections -  
 XX  
 XX Claim 26; Fig 6A-B; 170pp; English.  
 XX  
 XX The present sequence represents a human interleukin-21 (IL-21)  
 XX protein. The specification also describes IL-22 polynucleotides and  
 XX polypeptides. The IL-21 polynucleotide was isolated from a cDNA library  
 XX of apoptotic T-cells. IL-21 and IL-22 may be useful in treating  
 XX deficiencies or disorders of the immune system, by activating or

CC inhibiting the proliferation, differentiation, or mobilization  
 CC (chemotaxis) of immune cells, treating or detecting deficiencies or  
 CC disorders of haematopoietic cells, to modulate haemostatic or  
 CC thrombolytic activity, in treating or detecting autoimmune disorders,  
 CC treating asthma (particularly allergic asthma) or other respiratory  
 CC problems, to treat and/or prevent organ rejection or graft-versus-host  
 CC disease (GVHD), to modulate inflammation, to treat or detect  
 CC hyperproliferative disorders, to treat or detect infectious agents, to  
 CC differentiate, proliferate and attract cells, leading to the  
 CC regeneration of tissues, IL-21 and IL-22 may also increase or decrease  
 CC the differentiation or proliferation of embryonic stem cells and  
 CC haematopoietic lineage, may be used to modulate mammalian  
 CC characteristics.  
 XX  
 SQ Sequence 197 AA;

Query Match 4.4%; Score 7; DB 21; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PVGCTCV 128  
 |||||  
 Db 186 pvgctcv 192

RESULT 49  
 AAU04951

ID AAU04951 standard; Protein; 197 AA.  
 AC AAU04951;  
 DT 24-OCT-2001 (first entry)  
 XX Human Interleukin 17C ligand, IL-17C.  
 KW Human; Interleukin-17C ligand; IL-17C; agonist; antagonist;  
 KW PRO1122; DNA 62377-1381-1; systemic lupus erythematosus;  
 KW rheumatoid arthritis; osteoarthritis; diabetes mellitus;  
 KW allergic disease; asthma; demyelinating disease;  
 KW degenerative cartilaginous disorder; transplantation associated disease.  
 XX Homo sapiens.  
 XX

Key Location/Qualifiers  
 FT Peptide 1..18  
 FT /label= Signal\_peptide  
 FT Region 3..25  
 FT /note= "Leucine zipper pattern"  
 FT Protein 19..197  
 FT /label= Mature\_IL\_17C  
 FT Region 32..38  
 FT /note= "N-myristoylation site"  
 FT Region 55..61  
 FT /note= "N-myristoylation site"  
 FT Region 99..125  
 FT /note= "Region homologous to IL-17"  
 FT Region 112..121  
 FT /note= "Tyrosine kinase phosphorylation site"  
 FT Region 133..139  
 FT /note= "N-myristoylation site"  
 XX  
 XX WO200146420-A2.  
 XX  
 XX 28-JUN-2001.  
 XX  
 XX 20-DEC-2000; 2000WO-US34956.  
 XX 23-DEC-1999; 99US-0172096.  
 XX 30-DEC-1999; 99WO-US31274.  
 XX 11-JAN-2000; 2000US-0175481.  
 XX 18-FEB-2000; 2000WO-US04341.  
 XX 02-MAR-2000; 2000WO-US05841.

PR 21-MAR-2000; 2000US-0191007.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 22-JUN-2000; 2000US-0213087.  
 PR 22-AUG-2000; 2000US-0644848.  
 PR 24-AUG-2000; 2000WO-US23328.  
 PR 24-OCT-2000; 2000US-0242837.  
 PR 10-NOV-2000; 2000WO-US30873.  
 PR 28-NOV-2000; 2000US-0253646.  
 PR 01-DEC-2000; 2000WO-US32678.  
 XX (GEPH ) GENENTECH INC.  
 XX  
 XX Chen J, Filvaroff E, Fong S, Goddard A, Godowski PJ, Grimaldi CJ,  
 PI Gurney AL, Li H, Hillan KJ, Tumas D, Van Lookeren M, Vandlen RL;  
 PI Watanabe CK, Williams PM, Wood WI, Yansura DG;  
 XX  
 DR WPI; 2001-451708/48.  
 DR N-PSDB; AAS09510.  
 XX  
 XX Novel PRO polypeptides homologous to interleukin-17, useful for the  
 PT diagnosis and treatment of immune related disease e.g. rheumatoid  
 PT arthritis and diabetes -  
 XX  
 PS Claim 10; Fig 4; 188pp; English.  
 XX  
 XX The sequence is PRO1122 which is the human Interleukin 17C ligand,  
 CC IL-17C, encoded by DNA 62377-1381-1. A composition  
 CC containing ant/agonists to the PRO polypeptides or individual components  
 CC are useful for treating a mammal with an immune related disease, e.g.  
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
 CC thrombocytopaenia, thyroiditis, diabetes mellitus, immune-mediated renal  
 CC disease, a demyelinating disease, an autoimmune or immune-mediated skin  
 CC disease, contact dermatitis, an allergic disease e.g. food  
 CC hypersensitivity, asthma, a transplantation associated disease, or a  
 CC chronic inflammatory demyelinating polyneuropathy. Treating a  
 CC degenerative cartilaginous disorder comprises administering a PRO1031 or  
 CC PRO1122 polypeptide agonist, or antagonist to the mammal. Numerous  
 CC examples of the diseases and disorders are given in the specification.  
 XX  
 SQ Sequence 197 AA;

Query Match 4.4%; Score 7; DB 22; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PVGCTCV 128  
 |||||  
 Db 186 pvgctcv 192

RESULT 50  
 AAU44485

ID AAU44485 standard; Protein; 206 AA.  
 XX  
 AC AAU44485;  
 XX  
 XX 27-MAR-2000 (first entry)  
 DT  
 XX Human Interleukin 17C with C-terminal Gly(His)8 tag, IL-17C.his.  
 XX  
 XX Interleukin; IL-17C.his; PRO1122 polypeptide; clone DNA62377-1381-1;  
 KW immunoprecipitation; IL-17 receptor extracellular domain; IL-17R ECD;  
 KW cytokine IL-17; hybridisation probe; antagonist; Gly(His)8 tag; agonist;  
 KW degenerative cartilaginous disorder; diagnose; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers

FT Peptide 1..18  
FT /label= Signal\_peptide  
FT Protein 19..197  
FT /label= Mature\_IL-17C\_polypeptide  
FT /note= "Used to treat degenerative cartilaginous  
FT disorder"  
FT Misc-difference 109  
FT /note= "Conserved Trp residue"  
FT Misc-difference 129  
FT /note= "Conserved Cys residue"  
FT Misc-difference 134  
FT /note= "Conserved Cys residue"  
FT Misc-difference 163  
FT /note= "Conserved Cys residue"  
FT Misc-difference 189  
FT /note= "Conserved Cys residue"  
FT Misc-difference 191  
FT /note= "Conserved Cys residue"  
FT Misc-difference 198..206  
FT /note= "C-terminal Gly(His)8 tag"

XX  
PN WO9960127-A2.

XX  
PD 25-NOV-1999.

XX  
PF 14-MAY-1999; 99WO-US10733.

XX  
PR 15-MAY-1998; 98US-0085579.

XX  
PR 23-DEC-1998; 98US-0113621.

XX  
PA (GETH ) GENENTECH INC.

XX  
PI Chen J, Filvaroff E, Goddard A, Gurney AL, Li H, Wood WI;

XX  
DR WPI; 2000-116314/10.

XX  
PT New polypeptides designated PRO1031 and PRO1122 used to treat a

PT degenerative cartilaginous disorder -

XX  
PS Example 11; Page 138-139; 141pp; English.

XX  
CC The present sequence is the human PRO1122 polypeptide, with a C-terminal  
CC Gly(His)8 tag, IL-17C.his, derived from the clone DNA62377-1381-1.  
CC This sequence is used in a competitive binding experiment for the  
CC immunoprecipitation of IL-17 receptor extracellular domain (IL-17R ECD).  
CC The entire coding region of IL-17C can be used as hybridisation probe.  
CC The PRO1122 polypeptide, agonist or antagonist, is used to diagnose and  
CC treat a degenerative cartilaginous disorder.

XX  
SQ Sequence 206 AA;

Query Match 4.4%; Score 7; DB 21; Length 206;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PVGCTCV 128

Db 186 pvgctcv 192

Search completed: April 14, 2002, 12:42:15  
Job time: 217 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protéin - protein search, using sw model

Run on: April 14, 2002, 12:40:53 ; Search time 14.48 Seconds  
(without alignments)  
841.708 Million cell updates/sec

Title: US-09-731-816-4  
Perfect score: 160  
Sequence: 1 NSARARAVLSAPHTLQLGP.....SIDKQAKLLIGPNDAPAG 160

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 6

Total number of hits satisfying chosen parameters: 416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	5.0	261	2 E83150	hypothetical prote
2	8	5.0	804	2 T14762	hypothetical prote
3	8	5.0	837	2 T19825	hypothetical prote
4	7	4.4	56	2 H82806	hypothetical prote
5	7	4.4	61	2 B64218	ribosomal protein
6	7	4.4	61	2 S62805	ribosomal protein
7	7	4.4	150	2 B72588	hypothetical prote
8	7	4.4	154	2 D69338	conserved hypotet
9	7	4.4	170	2 H71805	hypothetical prote
10	7	4.4	170	2 B64713	hypothetical prote
11	7	4.4	209	2 T44553	conserved hypotet
12	7	4.4	243	2 A70623	probable lpgu prot
13	7	4.4	288	2 T35695	probable transcrip
14	7	4.4	293	2 H70625	hypothetical prote
15	7	4.4	307	1 G64717	pyridoxal phosphat
16	7	4.4	307	2 D71801	pyridoxal phosphat
17	7	4.4	332	2 T35917	probable regulator
18	7	4.4	348	1 S17779	stri protein - Str
19	7	4.4	360	2 F86775	hypothetical prote
20	7	4.4	395	2 T19144	hypothetical prote
21	7	4.4	399	2 T32126	hypothetical prote
22	7	4.4	467	2 D81051	multidrug efflux p
23	7	4.4	467	2 F81825	probable outer mem
24	7	4.4	497	2 I55514	Lic-2 - rat
25	7	4.4	946	2 S27921	nuclear antigen EB
26	7	4.4	972	2 A70619	excinuclease ABC c
27	7	4.4	1088	2 T18559	hypothetical prote
28	7	4.4	1490	2 T24502	hypothetical prote
29	7	4.4	1822	2 T14106	probable GTPase-ac

DNA replication/re  
MHC class II histo  
hypothetical prote  
brain neuron cytop  
hypothetical prote  
hypothetical prote  
capsid assembly pr  
hypothetical prote  
hypothetical prote  
probable membrane  
probable integral  
hypothetical prote  
hypothetical prote  
chorion class A pr  
hypothetical prote  
outer membrane pro  
14K cell fusion pr  
acyl carrier prote  
hypothetical prote  
mobB protein - Esc  
cytotoxic T-lympho  
hypothetical prote  
cytotoxic T-lympho  
immediate-early pr  
hypothetical prote  
carboxylesterase (  
conserved hypotet  
probable peptide m  
hypothetical prote  
high mobility grou  
superoxide dismuta  
hypothetical prote  
hypothetical prote  
probable N-6 adeni  
3-methyladenine DN  
conserved hypotet  
hypothetical prote  
probable mHf prot  
hypothetical prote  
ubiquitin--protein  
G/U mismatch-speci  
peroxinectin-like  
phospholipase C (E  
conserved hypotet  
ribosomal protein  
amidotransferase H  
hypothetical DNaj  
hypothetical prote  
hypothetical prote  
hypothetical prote  
probable transcrip  
redz protein - Str  
hypothetical prote  
achaete-scute homo  
hypothetical prote  
early ELA 25K prot  
Cdb1 protein - she  
hypothetical prote  
hypothetical prote  
probable biotin--a  
hypothetical prote  
ribosomal protein  
probable ssDNA-bin  
4-hydroxy-2-oxoval  
probable ABC trans  
abc transporter AT  
ornithine carbamoy  
probable hydrolase  
sporulation protei  
tail fiber protein

103	6	3.8	268	2	T24795	hypothetical prote	176	6	3.8	376	2	S57867	oncogene 1 - human
104	6	3.8	269	2	I54432	MHC class II histo	177	6	3.8	380	2	A69277	Na+/H+ antiporter
105	6	3.8	269	2	T00088	ABC-type transport	178	6	3.8	380	2	T00937	hypothetical prote
106	6	3.8	271	2	E85813	hypothetical prote	179	6	3.8	381	1	S15809	cytochrome P450 Cy
107	6	3.8	271	2	G64959	hypothetical prote	180	6	3.8	381	1	CMBO	chymosin (EC 3.4.2
108	6	3.8	272	2	E83363	hypothetical prote	181	6	3.8	383	2	G83785	glucose-1-phosphat
109	6	3.8	275	2	D70737	hypothetical prote	182	6	3.8	383	2	A55739	(MIC) protein MHC
110	6	3.8	277	2	T45332	hypothetical prote	183	6	3.8	383	2	T25865	hypothetical prote
111	6	3.8	282	2	T20977	hypothetical prote	184	6	3.8	387	2	JC2032	regulatory protein
112	6	3.8	283	2	F75265	conserved hypotet	185	6	3.8	388	2	G83778	hypothetical prote
113	6	3.8	283	2	T27423	hypothetical prote	186	6	3.8	396	2	G69504	conserved hypotet
114	6	3.8	283	2	S27859	piroplasm surface	187	6	3.8	398	2	F75254	acetate kinase - D
115	6	3.8	283	2	T24794	hypothetical prote	188	6	3.8	402	2	F72340	probable carboxyl-
116	6	3.8	286	2	E83048	hypothetical prote	189	6	3.8	403	2	I54192	aminomethyltransfe
117	6	3.8	288	2	C56281	7alpha-cephem-meth	190	6	3.8	404	2	A35407	tryptophan synthas
118	6	3.8	290	2	A70208	conserved hypotet	191	6	3.8	404	2	C81274	hypothetical prote
119	6	3.8	291	2	E84408	hypothetical prote	192	6	3.8	405	2	F81273	hypothetical prote
120	6	3.8	291	2	T50869	hypothetical prote	193	6	3.8	406	2	S76451	hypothetical prote
121	6	3.8	294	2	B70975	hypothetical prote	194	6	3.8	408	2	T44859	glycosyltransferas
122	6	3.8	295	2	T40654	probable signal tr	195	6	3.8	408	2	G81273	hypothetical prote
123	6	3.8	301	2	S73023	hypothetical prote	196	6	3.8	408	2	S66705	hypothetical prote
124	6	3.8	302	1	D64814	ybhK protein - Esc	197	6	3.8	409	2	H82842	DNA/pantothenate m
125	6	3.8	302	2	C85586	probable structura	198	6	3.8	412	2	B44418	surface antigen -
126	6	3.8	304	2	F75278	conserved hypotet	199	6	3.8	414	2	C72036	conserved hypotet
127	6	3.8	305	2	T47844	hypothetical prote	200	6	3.8	414	2	F86587	CT605 hypothetical
128	6	3.8	306	2	T21938	hypothetical prote	201	6	3.8	415	2	T13651	hypothetical prote
129	6	3.8	307	2	I40542	xpsR protein - Pse	202	6	3.8	417	1	S47793	valine--pyruvate t
130	6	3.8	310	2	A81298	formate dehydrogen	203	6	3.8	417	2	E86032	hypothetical prote
131	6	3.8	311	1	H71720	protein-export mem	204	6	3.8	420	2	T36072	hypothetical prote
132	6	3.8	316	2	S16681	homeotic protein -	205	6	3.8	421	2	B81864	probable glutamate
133	6	3.8	318	2	A84183	methanol dehydroge	206	6	3.8	421	2	B81079	glutamate dehydrog
134	6	3.8	321	1	BVECBF	birA bifunctional	207	6	3.8	423	2	T07150	G-box binding fact
135	6	3.8	321	2	A86089	hypothetical prote	208	6	3.8	426	2	C75297	probable protein s
136	6	3.8	321	2	T04720	hypothetical prote	209	6	3.8	427	1	VHVNPV	nucleoprotein - pi
137	6	3.8	322	2	T38399	probable amidohydr	210	6	3.8	427	2	T39430	mitochondrial impo
138	6	3.8	323	2	T13005	hypothetical prote	211	6	3.8	434	2	B85754	probable polysacch
139	6	3.8	323	2	C83282	hypothetical prote	212	6	3.8	441	2	T35083	histidinol dehydro
140	6	3.8	326	1	E69690	transcription repr	213	6	3.8	443	2	B81122	phage virion prote
141	6	3.8	327	2	C85749	hypothetical prote	214	6	3.8	445	2	D83819	hypothetical prote
142	6	3.8	327	2	A64884	probable transport	215	6	3.8	447	2	S59742	Slf1 protein - Yea
143	6	3.8	328	2	S73842	MG350 homolog G12-	216	6	3.8	448	2	T25552	hypothetical prote
144	6	3.8	328	2	E86880	hypothetical prote	217	6	3.8	451	2	T05579	hypothetical prote
145	6	3.8	330	2	C96611	hypothetical prote	218	6	3.8	452	1	WHRTF	phenylalanine 4-mo
146	6	3.8	331	1	A54932	zeta-crystallin /	219	6	3.8	453	1	WHRTF	phenylalanine 4-mo
147	6	3.8	331	2	F84823	probable peroxisom	220	6	3.8	453	2	S15758	phenylalanine 4-mo
148	6	3.8	336	2	S69524	hypothetical prote	221	6	3.8	453	2	F82702	conserved hypotet
149	6	3.8	336	2	T38175	hypothetical prote	222	6	3.8	455	2	T12041	cysteine proteinas
150	6	3.8	340	2	A83401	hypothetical prote	223	6	3.8	459	2	G86212	hypothetical prote
151	6	3.8	341	2	T35027	hypothetical prote	224	6	3.8	462	1	TVMSMC	transforming prote
152	6	3.8	341	2	C83067	hypothetical prote	225	6	3.8	462	1	TVMSM2	transforming prote
153	6	3.8	343	2	C75260	conserved hypotet	226	6	3.8	463	2	T21042	hypothetical prote
154	6	3.8	344	2	G70726	DNA-binding protei	227	6	3.8	464	2	F72512	hypothetical prote
155	6	3.8	345	2	T29261	hypothetical prote	228	6	3.8	465	2	T34690	hypothetical prote
156	6	3.8	346	1	TVHMS	protein kinase (EC	229	6	3.8	466	2	T32478	hypothetical prote
157	6	3.8	346	1	TVMRMS	protein kinase (EC	230	6	3.8	469	2	C82495	glycerol-3-phospha
158	6	3.8	348	2	C69733	PBSX prophage ORF	231	6	3.8	474	2	B69494	phenylalanyl-tRNA
159	6	3.8	348	2	G69948	phage-related prot	232	6	3.8	475	2	T01352	hypothetical prote
160	6	3.8	348	2	I49262	cyclin cdk inhibit	233	6	3.8	479	2	S21567	translation elonga
161	6	3.8	351	2	T50594	probable oxidoredu	234	6	3.8	480	2	F81965	lactaldehyde dehyd
162	6	3.8	351	2	T19372	hypothetical prote	235	6	3.8	480	2	A81023	aldehyde dehydroge
163	6	3.8	353	2	H82644	NADP-alcohol dehyd	236	6	3.8	483	2	C82587	cationic amino aci
164	6	3.8	354	2	C83577	hypothetical prote	237	6	3.8	492	2	T02458	hypothetical prote
165	6	3.8	358	2	D83374	hypothetical prote	238	6	3.8	493	2	T18789	hypothetical prote
166	6	3.8	360	2	S48566	hypothetical prote	239	6	3.8	495	2	G70593	probable cationic
167	6	3.8	360	2	E70892	hypothetical prote	240	6	3.8	496	2	A37979	cartilage matrix p
168	6	3.8	361	2	D72721	hypothetical prote	241	6	3.8	497	2	F71472	cysteine--tRNA lig
169	6	3.8	363	2	T36024	conserved hypotet	242	6	3.8	497	2	T51907	hypothetical prote
170	6	3.8	365	2	T07151	G-box binding fact	243	6	3.8	502	2	D70806	probable fadD17 pr
171	6	3.8	366	2	G64449	modification methy	244	6	3.8	504	2	T01603	hypothetical prote
172	6	3.8	369	2	S72734	DNA-binding protei	245	6	3.8	507	2	H69186	conserved hypotet
173	6	3.8	369	2	G82829	membrane fusion pr	246	6	3.8	509	2	JA0148	photosystem II chl
174	6	3.8	371	2	H83285	probable two-compo	247	6	3.8	509	2	T06855	photosystem II chl
175	6	3.8	376	2	T48245	hypothetical prote	248	6	3.8	510	2	G72464	hypothetical prote

249	6	3.8	512	2	C81105	peptidyl-prolyl ci	322	6	3.8	776	2	S44784	C30C11.4 protein -
250	6	3.8	512	2	H81909	probable peptidylp	323	6	3.8	780	2	T21708	hypothetical prote
251	6	3.8	515	2	JC5458	inulinase (EC 3.2.	324	6	3.8	781	2	A69312	DNA polymerase B1
252	6	3.8	515	3	JC7533	inulinase (EC 3.2.	325	6	3.8	786	2	S22155	oncogene 1 (tre-2
253	6	3.8	516	2	JE0301	inulinase (EC 3.2.	326	6	3.8	790	2	A35797	probable DNA-bindi
254	6	3.8	517	2	T35295	probable aminopept	327	6	3.8	806	2	S22765	heterogeneous ribo
255	6	3.8	518	2	S61920	B-alpha pheromone-	328	6	3.8	836	2	S61570	regulatory protein
256	6	3.8	521	2	JC5220	protein-tyrosine-p	329	6	3.8	838	2	S28911	gene DN10 protein
257	6	3.8	532	2	A57173	oculocutaneous alb	330	6	3.8	862	1	A49346	aldehyde dehydroge
258	6	3.8	534	2	T33393	hypothetical prote	331	6	3.8	867	1	GNLJSA	pol polyprotein -
259	6	3.8	540	2	T03309	probable terminase	332	6	3.8	867	1	GNLJMP	pol polyprotein (c
260	6	3.8	543	2	C84970	flagellar hook-ass	333	6	3.8	870	2	T01310	hypothetical prote
261	6	3.8	545	2	T27614	hypothetical prote	334	6	3.8	878	2	B71460	probable outer mem
262	6	3.8	548	2	T05671	hypothetical prote	335	6	3.8	879	1	JDVLC	DNA-directed DNA p
263	6	3.8	550	2	T36746	probable serine/th	336	6	3.8	886	2	F83962	penicillin-binding
264	6	3.8	550	2	E85574	probable fumarate	337	6	3.8	887	1	DEGCPV	pyruvate dehydroge
265	6	3.8	556	2	A32466	numb protein - fru	338	6	3.8	887	2	B84954	pyruvate dehydroge
266	6	3.8	560	2	C86716	2-oxoglutarate dec	339	6	3.8	887	2	F85494	hypothetical prote
267	6	3.8	562	2	B70609	hypothetical prote	340	6	3.8	891	1	JN0867	peroxinectin-like
268	6	3.8	566	2	T30362	immediate early tr	341	6	3.8	894	2	PN0667	probable finger pr
269	6	3.8	568	2	H64879	probable membrane	342	6	3.8	894	2	S51245	stable tubule only
270	6	3.8	576	1	B70558	probable ABC trans	343	6	3.8	906	2	JC5963	pyruvate dehydroge
271	6	3.8	578	2	T41715	hypothetical prote	344	6	3.8	978	2	C59237	hypothetical prote
272	6	3.8	579	2	T47705	hypothetical prote	345	6	3.8	980	2	T49570	hypothetical prote
273	6	3.8	581	2	T38864	probable regulator	346	6	3.8	992	2	G70950	mitotic control pr
274	6	3.8	583	2	T12576	probable phosphate	347	6	3.8	1000	2	HE0110	type III restricti
275	6	3.8	584	2	H86531	CT085 hypothetical	348	6	3.8	1001	2	H64593	probable ABC-type
276	6	3.8	584	2	T20992	conserved hypothet	349	6	3.8	1011	2	T07712	hypothetical prote
277	6	3.8	586	2	T25695	nitrate-inducible	350	6	3.8	1018	2	T23318	hypothetical prote
278	6	3.8	590	2	A45772	probable acyl-CoA	351	6	3.8	1020	2	A29355	fibronectin - chic
279	6	3.8	592	2	H83580	Shb protein - huma	352	6	3.8	1021	2	T23252	hypothetical prote
280	6	3.8	596	2	I38228	hypothetical prote	353	6	3.8	1025	2	T18376	multidrug resistan
281	6	3.8	599	2	T24333	hypothetical prote	354	6	3.8	1026	1	TLBP74	tail fiber protein
282	6	3.8	607	2	H84514	hypothetical prote	355	6	3.8	1037	2	E81980	pillus-associated p
283	6	3.8	608	2	H71379	probable tpr prote	356	6	3.8	1050	2	S57488	fimbrial protein p
284	6	3.8	609	2	T00904	hypothetical prote	357	6	3.8	1071	2	S38164	ATP-binding protei
285	6	3.8	610	2	T16761	hypothetical prote	358	6	3.8	1077	2	T01474	hypothetical prote
286	6	3.8	611	2	T44560	hypothetical prote	359	6	3.8	1107	1	S52517	myosin I heavy cha
287	6	3.8	616	1	I64056	secretion protein	360	6	3.8	1109	2	S53601	myosin-IC [similar
288	6	3.8	616	2	J01441	hypothetical 67K p	361	6	3.8	1134	2	G69269	polydoprotein-bind
289	6	3.8	622	2	A70414	NADH dehydrogenase	362	6	3.8	1138	2	A48944	parasporal crystal
290	6	3.8	625	2	T30022	hypothetical prote	363	6	3.8	1139	2	B70954	hypothetical prote
291	6	3.8	626	2	F71143	hypothetical prote	364	6	3.8	1154	2	S69206	regulator protein
292	6	3.8	627	1	A43300	squalene-hopene cy	365	6	3.8	1194	2	T37503	probable chromosom
293	6	3.8	632	1	VGNSY	surface glycoprote	366	6	3.8	1202	2	T37867	hypothetical prote
294	6	3.8	638	2	T36309	probable transcrip	367	6	3.8	1203	2	S26650	DNA-binding protei
295	6	3.8	639	2	A32545	protein kinase C (	368	6	3.8	1215	2	H84513	probable disease r
296	6	3.8	640	2	T46936	hypothetical prote	369	6	3.8	1228	2	S60085	nitrate reductase
297	6	3.8	644	2	B70420	NADH dehydrogenase	370	6	3.8	1237	2	T08608	hypothetical prote
298	6	3.8	646	2	H82555	c-type cytochrome	371	6	3.8	1246	2	T00826	hypothetical prote
299	6	3.8	646	2	E82133	ATP-dependent heli	372	6	3.8	1261	2	T50065	hypothetical prote
300	6	3.8	649	2	B96729	hypothetical prote	373	6	3.8	1274	2	T16251	hypothetical prote
301	6	3.8	651	2	T42644	hypothetical prote	374	6	3.8	1324	2	S51622	cut3 protein - fis
302	6	3.8	652	1	S52695	2',3'-cyclic-nucle	375	6	3.8	1331	2	T04938	hypothetical prote
303	6	3.8	657	2	B84791	hypothetical prote	376	6	3.8	1360	2	T34302	cell polarity prot
304	6	3.8	659	2	D84633	probable multispin	377	6	3.8	1407	2	S59823	probable membrane
305	6	3.8	675	2	T35959	methylmalonyl-CoA	378	6	3.8	1433	2	T30261	chitinase (EC 3.2.
306	6	3.8	676	2	S67136	hypothetical prote	379	6	3.8	1445	2	T10728	probable gag/pol p
307	6	3.8	688	2	T33709	hypothetical prote	380	6	3.8	1446	2	T04528	myosin heavy chain
308	6	3.8	708	2	A47176	probable transmemb	381	6	3.8	1452	1	S17669	protein-tyrosine-p
309	6	3.8	716	1	T01507	replication licens	382	6	3.8	1452	1	S17670	protein-tyrosine-p
310	6	3.8	719	2	T52510	hypothetical prote	383	6	3.8	1464	1	S29159	glutamate receptor
311	6	3.8	728	2	B81311	probable bacteriop	384	6	3.8	1464	2	A43274	N-methyl D-asparta
312	6	3.8	729	2	T10692	polynucleotide ade	385	6	3.8	1464	2	S47555	N-methyl D-asparta
313	6	3.8	730	2	T00478	probable translati	386	6	3.8	1479	2	T17401	transcription regu
314	6	3.8	739	2	S32948	hypF protein - Rho	387	6	3.8	1502	2	T14278	myosin-like protei
315	6	3.8	744	2	T13429	hypothetical prote	388	6	3.8	1515	2	S51824	myosin heavy chain
316	6	3.8	755	2	JN0648	hydrogenase expres	389	6	3.8	1551	1	A43364	M polyprotein prec
317	6	3.8	755	2	S23441	hypothetical prote	390	6	3.8	1611	2	T38236	hypothetical prote
318	6	3.8	765	2	C75100	hypothetical prote	391	6	3.8	1628	2	T38055	hypothetical prote
319	6	3.8	768	2	T35465	hypothetical plasm	392	6	3.8	1655	2	T32633	hypothetical prote
320	6	3.8	773	2	JU0135	aldehyde dehydroge	393	6	3.8	1681	2	S59693	hypothetical prote
321	6	3.8	775	1	WMBE19	ribonucleoside-dip	394	6	3.8	1733	2	D70887	probable polyketid

395 6 3.8 1796 2 S65004 probable membrane  
396 6 3.8 1797 2 T21889 hypothetical prote  
397 6 3.8 1805 2 T21888 hypothetical prote  
398 6 3.8 1816 2 A84845 probable ABC trans  
399 6 3.8 1859 2 S63325 probable membrane  
400 6 3.8 1859 2 S64633 probable membrane  
401 6 3.8 2115 2 S38480 nonstructural prot  
402 6 3.8 2205 1 MNWVRN nonstructural poly  
403 6 3.8 2342 2 T18200 fatty-acid synthas  
404 6 3.8 2442 2 T08621 centrosome associa  
405 6 3.8 2479 1 MNWVRA nonstructural poly  
406 6 3.8 2481 2 A43908 fibronectin - Afri  
407 6 3.8 2523 2 F70846 probable PPE prote  
408 6 3.8 2560 1 I40457 peptide synthetase  
409 6 3.8 2895 2 T08437 hyperplastic discs  
410 6 3.8 3149 1 Q08B8 Bpfl protein - hu  
411 6 3.8 3455 2 B82519 hemagglutinin-like  
412 6 3.8 3573 2 S23070 erythronolide synt  
413 6 3.8 3587 2 I40486 surfactin syntheta  
414 6 3.8 3598 2 I40485 surfactin syntheta  
415 6 3.8 4427 2 PN0637 polyketide synthas  
416 6 3.8 4688 2 F82885 hypothetical prote

## ALIGNMENTS

RESULT 1  
E83150  
hypothetical protein PA3964 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
R:Accession: E83150  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: AB2950; MUID:20437337  
A:Accession: E83150  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-261 <STO>  
A:Cross-references: GB:AE004814; GB:AE004091; NID:g9950147; PIDN:AAG07351.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3964

Query Match 5.0%; Score 8; DB 2; Length 261;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 AGGRPADR 38  
|||||  
Db 80 AGGRPADR 87

RESULT 2  
T14762  
hypothetical protein DKFZp434A014.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
R:Accession: T14762  
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, August 1999  
A:Reference number: Z18181  
A:Accession: T14762  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-804 <WAM>  
A:Cross-references: EMBL:AL110224  
A:Experimental source: adult testis; clone DKFZp434A014

C:Genetics:  
A:Note: DKFZp434A014.1

Query Match 5.0%; Score 8; DB 2; Length 804;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LFGEEDVR 88  
|||||  
Db 663 LFGEEDVR 670

## RESULT 3

T19825  
hypothetical protein C38D4.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Aug-2000  
C:Accession: T19825  
R:Colles, L.  
submitted to the EMBL Data Library, October 1994  
A:Reference number: Z19183  
A:Accession: T19825  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-837 <WIL>  
A:Cross-references: EMBL:Z46241; PIDN:CAA86318.1; GSPDB:GN00021; CESP:C38D4.5  
A:Experimental source: clone C38D4  
C:Genetics:  
A:Gene: CESP:C38D4.5  
A:Map position: 3  
A:Introns: 31/3; 81/2; 137/3; 277/2; 435/3; 507/3; 544/2; 604/1; 646/2; 707/1; 791/3  
R:Superfamily: WW repeat homology  
C:Superfamily: WW repeat homology <WWR>  
F:96-133/Domain: WW repeat homology <WWR>

Query Match 5.0%; Score 8; DB 2; Length 837;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 PTNLRVS 50  
|||||  
Db 312 PTNLRVS 319

## RESULT 4

H82806  
hypothetical protein XF0431 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
R:Accession: H82806  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: AB2515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82806  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-56 <SIM>  
A:Cross-references: GB:AE003894; GB:AE003849; NID:g9105267; PIDN:AAF83241.1; GSPDB:GN  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
., F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv

M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
 A;Reference number: A59328  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: XF0431

Query Match 4.4%; Score 7; DB 2; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 86 DVFRSA 92  
 |||||  
 Db 31 DVFRSA 37

RESULT 5  
 B64218  
 Ribosomal protein S14 - Mycoplasma genitalium  
 C;Species: Mycoplasma genitalium  
 C;Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 07-Dec-1999  
 C;Accession: B64218  
 R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
 M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.  
 ; C.A.; Venter, J.C.  
 Science 270, 397-403, 1995  
 A;Title: The minimal gene complement of Mycoplasma genitalium.  
 A;Reference number: A64200; MUID:96026346  
 A;Accession: B64218  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-61 <TIGR>  
 A;Cross-references: GB:U39695; GB:I43967; NID:gl045833; PID:gl045848; TIGR:MG164  
 A;Experimental source: strain G-37  
 C;Genetics:  
 A;Genetic code: SGC3  
 C;Superfamily: Escherichia coli ribosomal protein S14

Query Match 4.4%; Score 7; DB 2; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 5.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 RARAVLS 10  
 |||||  
 Db 29 RARAVLS 35

RESULT 6  
 S62805  
 ribosomal protein S14 - Mycoplasma pneumoniae (strain ATCC 29342)  
 N;Alternate names: hypothetical protein GT9\_Orf61  
 C;Species: Mycoplasma pneumoniae  
 A;Variety: ATCC 29342  
 C;Date: 23-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-Dec-1999  
 C;Accession: S62805; S73979  
 R;Hilbert, H.; Himmelreich, R.; Plagens, H.; Herrmann, R.  
 Nucleic Acids Res. 24, 628-639, 1996  
 A;Title: Sequence analysis of 56 Kb from the genome of the bacterium Mycoplasma pneumoniae  
 A;Reference number: S62797; MUID:96177562  
 A;Accession: S62805  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: DNA  
 A;Residues: 1-61 <HIL>  
 A;Cross-references: EMBL:U34795; NID:gl215683; PIDN:AAC43703.1; PID:gl215710  
 R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
 Nucleic Acids Res. 24, 4420-4449, 1996  
 A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
 A;Reference number: S73327; MUID:97105885  
 A;Accession: S73979  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-61 <HTM>

A;Cross-references: EMBL:AE000061; GB:U00089; NID:gl674336; PIDN:AAB96301.1; PID:gl616  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 199.  
 C;Genetics:  
 A;Gene: rpsN  
 A;Genetic code: SGC3  
 C;Superfamily: Escherichia coli ribosomal protein S14  
 C;Keywords: protein biosynthesis; ribosome

Query Match 4.4%; Score 7; DB 2; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 5.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 RARAVLS 10  
 |||||  
 Db 29 RARAVLS 35

RESULT 7  
 B72588  
 hypothetical protein APE1175 - Aeropyrum pernix (strain K1)  
 C;Species: Aeropyrum pernix  
 C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
 C;Accession: B72588  
 R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; T  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J  
 DNA Res. 6, 83-101, 1999  
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aer  
 A;Reference number: A72450; MUID:99310339  
 A;Accession: B72588  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-150 <KAW>  
 A;Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAAS0160.1; PID:dl043946; PID:  
 A;Experimental source: strain K1  
 C;Genetics:  
 A;Gene: APE1175  
 C;Superfamily: Aeropyrum pernix hypothetical protein APE1175

Query Match 4.4%; Score 7; DB 2; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 31 AGRRPAD 37  
 |||||  
 Db 27 AGRRPAD 33

RESULT 8  
 D69338  
 conserved hypothetical protein AF0708 - Archaeoglobus fulgidus  
 C;Species: Archaeoglobus fulgidus  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C;Accession: D69338  
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Do  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arc  
 A;Reference number: A69250; MUID:98049343  
 A;Accession: D69338  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-154 <KLE>  
 A;Cross-references: GB:AE001056; GB:AE000782; NID:g2689379; PIDN:AAB90533.1; PID:g26

Query Match 4.4%; Score 7; DB 2; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 31 AGRRPAD 37  
 |||||  
 Db 27 AGRRPAD 33

Qy 6 RAVLSAF 12  
 Db 129 RAVLSAF 135

## RESULT 9

H71805  
 hypothetical protein jhpl453 - Helicobacter pylori (strain J99)  
 C:Species: Helicobacter pylori  
 A:Variety: strain J99  
 C:date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
 C:Accession: H71805  
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
 Nature 397, 176-180, 1999  
 A:title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
 A:Reference number: A71800; MUID:99120557  
 A:Accession: H71805  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-170 <ARN>  
 A:Cross-references: GB:AE001567; GB:AE001439; NID:g4156065; PIDN:AAD07035.1; PID:g415608  
 A:Experimental source: strain J99  
 C:Genetics:  
 A:Gene: jhpl453

Query Match 4.4%; Score 7; DB 2; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 132 EKDADSI 138  
 Db 69 EKDADSI 75

## RESULT 10

B64713  
 hypothetical protein HPI546 - Helicobacter pylori (strain 26695)  
 C:Species: Helicobacter pylori  
 C:date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
 C:Accession: B64713  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.  
 Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
 A:title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:97394467  
 A:Accession: B64713  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-170 <TOM>  
 A:Cross-references: GB:AE000652; GB:AE000511; NID:g2314720; PIDN:AAD08590.1; PID:g231473

Query Match 4.4%; Score 7; DB 2; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 132 EKDADSI 138  
 Db 69 EKDADSI 75

## RESULT 11

T44553  
 conserved hypothetical protein PA0629 [imported] - Pseudomonas aeruginosa  
 C:Species: Pseudomonas aeruginosa  
 C:date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 31-Dec-2000  
 C:Accession: T44553; C83568  
 R:Nakayama, K.; Takashima, K.; Ishihara, H.; Shinomiya, T.; Kageyama, M.; Kanaya, S.; Oh

submitted to the EMBL Data Library, August 1999  
 A:Description: Genetic relationship between bacteriophages and bacteriophages.

A:Reference number: Z22790  
 A:Accession: T44553  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-209 <NAK>  
 A:Cross-references: EMBL:AB030825; PIDN:BAA83168.1  
 A:Experimental source: strain PA01  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
 Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: C83568  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-209 <STO>  
 A:Cross-references: GB:AE004498; GB:AE004091; NID:g9946491; PIDN:AAG04018.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA0629  
 A:Note: lys  
 C:Superfamily: Haemophilus influenzae hypothetical protein H1415

Query Match 4.4%; Score 7; DB 2; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RARAVLS 10  
 Db 203 RARAVLS 209

## RESULT 12

A70623  
 probable lpqU protein - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C:Accession: A70623  
 R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
 & Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
 A:Reference number: A70500; MUID:9825987  
 A:Accession: A70623  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-243 <COL>  
 A:Cross-references: GB:Z92539; GB:AL123456; NID:g3261714; PIDN:CAB06857.1; PID:e30455  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: lpqU

Query Match 4.4%; Score 7; DB 2; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 AGGRPAD 37  
 Db 51 AGGRPAD 57

## RESULT 13

T35695  
 probable transcription regulator - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 01-Dec-2000

C;Accession: T35695; T35132  
 R;Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, July 1998  
 A;Reference number: Z21587  
 A;Accession: T35695  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-288 <HAR>  
 A;Cross-references: EMBL:AL031031; PIDN:CAAL19863.1; GSPDB:GN00070; SCOEDB:SC7C7.17  
 A;Experimental source: strain A3(2)  
 R;Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, December 1997  
 A;Reference number: Z21569  
 A;Accession: T35132  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 260-288 <HA2>  
 A;Cross-references: EMBL:AL020958; PIDN:CAAL15868.1; GSPDB:GN00070; SCOEDB:SC4H8.01  
 A;Experimental source: strain A3(2)  
 C;Genetics:  
 A;Gene: SCOEDB:SC7C7.17; SCOEDB:SC4H8.01

Query Match 4.4%; Score 7; DB 2; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ARARAVL 9  
 |||||

Db 174 ARARAVL 180

## RESULT 14

H70625  
 hypothetical protein Rv1045 - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C;Accession: H70625  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, R.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295987

A;Accession: H70625  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA

A;Residues: 1-293 <COL>  
 A;Cross-references: GB:Z92539; GB:AL123456; NID:g3261714; PIDN:CAB06867.1; PID:g1869983  
 A;Experimental source: strain H37RV

C;Genetics:

A;Gene: Rv1045

C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1045

Query Match 4.4%; Score 7; DB 2; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 PPTNLR 48  
 |||||

Db 8 PPTNLR 14

## RESULT 15

G64717  
 pyridoxal phosphate biosynthetic protein A - Helicobacter pylori (strain 26695)  
 C;Species: Helicobacter pylori  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: G64717  
 R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKennd

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey  
 Nature 388, 539-547, 1997  
 A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser  
 A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A;Reference number: A64520; MUID:97394467  
 A;Accession: G64717  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-307 <TOM>  
 A;Cross-references: GB:AE000655; GB:AE000511; NID:g2314757; PIDN:AAD08621.1; PID:g23  
 C;Superfamily: pdxA protein

Query Match 4.4%; Score 7; DB 1; Length 307;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GLFGEED 86  
 |||||

Db 200 GLFGEED 206

## RESULT 16

D71801  
 pyridoxal phosphate biosynthetic protein A - Helicobacter pylori (strain J99)  
 C;Species: Helicobacter pylori  
 A;Variety: strain J99  
 C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 26-Aug-1999  
 C;Accession: D71801  
 R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D  
 ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.  
 Nature 397, 176-180, 1999  
 A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric  
 A;Reference number: A71800; MUID:99120557

A;Accession: D71801

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-307 <ARN>

A;Cross-references: GB:AE001570; GB:AE001439; NID:g4156108; PIDN:AAD07064.1; PID:g41

A;Experimental source: strain J99

C;Genetics:

A;Gene: pdxA

C;Superfamily: pdxA protein

Query Match 4.4%; Score 7; DB 2; Length 307;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GLFGEED 86  
 |||||

Db 200 GLFGEED 206

## RESULT 17

T35917  
 probable regulatory protein - Streptomyces coelicolor  
 C;Species: Streptomyces coelicolor  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
 C;Accession: T35917  
 R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream  
 submitted to the EMBL Data Library, January 1998

A;Reference number: Z21593

A;Accession: T35917

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-332 <SAU>

A;Cross-references: EMBL:AL035212; PIDN:CAA22801.1; GSPDB:GN00070; SCOEDB:SC9B2.21c

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SC9B2.21c

C;Superfamily: methanol dehydrogenase regulatory protein

Query Match 4.4%; Score 7; DB 2; Length 332;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ARARAVL 9  
| | | | |  
Db 266 ARARAVL 272

RESULT 18  
S1779  
stri protein - Streptomyces griseus  
C:Species: Streptomyces griseus  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
A:Reference number: S17779  
R:Nangouri, K.; Piepersberg, W.  
Mol. Gen. Genet. 228, 459-469, 1991  
A:Title: Genetics of streptomycin production in Streptomyces griseus: nucleotide sequence  
A:Reference number: S17775; MUID:91375432  
A:Accession: S17779  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-348 <MAN>  
A:Cross-references: GB:Y00459; GB:S55493; NID:g1621271; PID:g49268  
C:Genetics:  
A:Gene: stri  
C:Superfamily: Streptomyces griseus stri protein

Query Match 4.4%; Score 7; DB 1; Length 348;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ARAVLSA 11  
| | | | |  
Db 217 ARAVLSA 223

RESULT 19  
F86775  
hypothetical protein hsc [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 31-Mar-2001  
A:Accession: F86775  
R:Boletini, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlich  
Genome Res. in press, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium.  
A:Reference number: A86625  
A:Accession: F86775  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <STO>  
A:Cross-references: GB:AE005176; NID:g12724175; PIDN:AAK05304.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: hsc  
C:Superfamily: histidinol phosphate aminotransferase

Query Match 4.4%; Score 7; DB 2; Length 360;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 NLRSVSP 51  
| | | | |  
Db 6 NLRSVSP 12

RESULT 20  
T19144  
hypothetical protein C09G5.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: T19144  
R:Palmer, S.  
submitted to the EMBL Data Library, November 1994  
A:Reference number: Z19080  
A:Accession: T19144  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-395 <WIL>  
A:Cross-references: EMBL:Z46791; PIDN:CAA86759.1; GSPDB:GN00020; CESP:C09G5.7  
A:Experimental source: clone C09G5  
C:Genetics:  
A:Gene: CESP:C09G5.7  
A:Map position: 2  
A:Introns: 20/2; 57/3; 153/3; 300/1; 357/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein C54G4.2

Query Match 4.4%; Score 7; DB 2; Length 395;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 AYRISYD 59  
| | | | |  
Db 351 AYRISYD 357

RESULT 21  
T32126  
hypothetical protein R07C3.11 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
A:Accession: T32126  
R:Lamar, B.; Kramer, J.  
submitted to the EMBL Data Library, July 1997  
A:Description: The sequence of C. elegans cosmid R07C3.  
A:Reference number: Z21125  
A:Accession: T32126  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-399 <LAM>  
A:Cross-references: EMBL:AF016686; PIDN:AAB66240.1; GSPDB:GN00020; CESP:R07C3.11  
A:Experimental source: strain Bristol N2; clone R07C3  
C:Genetics:  
A:Gene: CESP:R07C3.11  
A:Map position: 2  
A:Introns: 363/2

Query Match 4.4%; Score 7; DB 2; Length 399;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 ISYDPA 62  
| | | | |  
Db 380 ISYDPA 386

RESULT 22  
D81051  
multidrug efflux pump channel protein NMB1714 [imported] - Neisseria meningitidis (st  
C:Species: Neisseria meningitidis  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
A:Accession: D81051  
R:rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, B.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.  
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;  
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A:Reference number: A81000; MUID:20175755  
A:Accession: D81051  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-467 <TET>  
A:Cross-references: GB:AE002521; GB:AE002098; NID:g7226962; PIDN:AAF42061.1; PID:g722696  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NM1714  
C:Superfamily: nodulation protein nodT

Query Match 4.4%; Score 7; DB 2; Length 467;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 REQARNA 27  
|||||||  
DB 235 REQARNA 241

RESULT 23  
F81825  
probable outer membrane lipoprotein NMA1968 [imported] - Neisseria meningitidis (strain  
C:Species: Neisseria meningitidis  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: F81825  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491.  
A:Reference number: A81775; MUID:20222556  
A:Accession: F81825  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-467 <PAR>  
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85188.1; PID:g738060  
A:Experimental source: serogroup A, strain 22491  
C:Genetics:  
A:Gene: mtrE; NMA1968  
C:Superfamily: nodulation protein nodT

Query Match 4.4%; Score 7; DB 2; Length 467;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 REQARNA 27  
|||||||  
DB 235 REQARNA 241

RESULT 24  
I55514  
LIC-2 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: I55514  
R:Hughes, S.M.; Vaughan, K.T.; Herskovits, J.S.; Vallee, R.B.  
J. Cell Sci. 108, 17-24, 1995  
A:Title: Molecular analysis of a cytoplasmic dynein light intermediate chain reveals hom  
A:Reference number: I55514; MUID:95256330  
A:Accession: I55514  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-497 <RES>  
A:Cross-references: EMBL:UI5138; NID:g619664; PIDN:AAA80334.1; PID:g619665

Query Match 4.4%; Score 7; DB 2; Length 497;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 KLLIGPN 154  
|||||||  
DB 9 KLLIGPN 15

RESULT 25  
S27921

nuclear antigen EBNA-3B - human herpesvirus 4  
C:Species: human herpesvirus 4, Epstein-Barr virus  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 26-Aug-1999  
C:Accession: S27921  
R:Sample, J.; Young, L.; Martin, B.; Chatman, T.; Kieff, E.; Rickinson, A.; Kieff, E.  
submitted to the EMBL Data Library, July 1990  
A:Reference number: S27920

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-946 <SAM>  
A:Cross-references: EMBL:M34440; NID:g330407; PIDN:AAA5894.1; PID:g330409  
C:Genetics:  
A:Introns: 119/3  
C:Superfamily: human herpesvirus 4 nuclear antigen EBNA-3B

Query Match 4.4%; Score 7; DB 2; Length 946;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 PTVVLR 103  
|||||||  
DB 450 PTVVLR 456

RESULT 26

A70619  
excinuclease ABC chain A - Mycobacterium tuberculosis (strain H37RV)  
N:Alternate names: uvra protein  
N:Contains: excision endonuclease ABC (EC 3.1.-.-) chain A  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 02-Feb-2001  
C:Accession: A70619  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordo  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd  
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete gen  
A:Reference number: A70500; MUID:98295987  
A:Accession: A70619  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-972 <COL>  
A:Cross-references: GB:Z85982; GB:AL123456; NID:g3261718; PIDN:CAB06633.1; PID:g1838  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: uvra

C:Superfamily: excinuclease ABC chain A; ATP-binding cassette homology  
C:Keywords: ATP; DNA binding; DNA repair; hydrolase; nucleotide binding; P-loop  
F:32-39/Region: nucleotide-binding motif A (P-loop)  
F:637-920/Domain: ATP-binding cassette homology <ABCE>  
F:654-661/Region: nucleotide-binding motif A (P-loop)

Query Match 4.4%; Score 7; DB 2; Length 972;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 LGPREQA 24  
|||||||  
DB 462 LGPREQA 468

RESULT 27

T18559  
hypothetical protein - Oxytricha fallax  
C:Species: Oxytricha fallax

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-May-2000  
C;Accession: T18559  
R;Seegmiller, A.; Williams, K.R.; Hammersmith, R.L.; Doak, T.G.; Witherspoon, D.; Messing, R.O. Biol. Evol. 13, 1351-1362, 1996  
A;Title: Internal eliminated sequences interrupting the Oxytricha 81 locus: allelic divergence  
A;Reference number: Z18972; MUID:97109822  
A;Accession: T18559  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-1088 <SEE>  
A;Cross-references: EMBL:U81495; NID:g2194111; PID:g2194112; PIDN:AAB61088.1  
A;Genetics:  
A;Genetic code: SGC5

Query Match 4.4%; Score 7; DB 2; Length 1088;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 SINSSID 143  
|||||||  
Db 1046 SINSSID 1052

RESULT 28  
T24502  
hypothetical protein T06D8.10 - Caenorhabditis elegans (fragment)  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C;Accession: T24502; T24585  
R;Swinburne, J.  
submitted to the EMBL Data Library, April 1995  
A;Reference number: Z19900  
A;Accession: T24502  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-1490 <WIL>  
A;Cross-references: EMBL:Z49129; PIDN:CAA88963.1; GSPDB:GN00020; CESP:T06D8.10  
A;Experimental source: clone T05B9  
R;Palmer, S.  
submitted to the EMBL Data Library, April 1995  
A;Reference number: Z19909  
A;Accession: T24585  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-1490 <WIL2>  
A;Cross-references: EMBL:Z49130; PIDN:CAA88973.1; GSPDB:GN00020; CESP:T06D8.10  
A;Experimental source: clone T06D8  
C;Genetics:  
A;Gene: CESP:T06D8.10  
A;Map position: 2  
A;Introns: 46/1; 123/2; 161/2; 222/3; 284/2; 340/1; 566/3; 627/3; 683/2; 774/3; 1400/2;

Query Match 4.4%; Score 7; DB 2; Length 1490;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 LGPREQA 24  
|||||||  
Db 305 LGPREQA 311

RESULT 29  
T14106  
probable GTPase-activating protein SPA-1 - rat  
N;Alternate names: protein p1294  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C;Accession: T14106  
R;Takeuchi, M.; Ide, N.; Hata, Y.; Takai, Y.  
submitted to the EMBL Data Library, September 1997  
A;Description: SPA-1 like protein identified through yeast two-hybrid screening using th

A;Reference number: Z17877  
A;Accession: T14106  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-1822 <YAK>  
A;Cross-references: EMBL:AF026504; NID:g2555182; PID:g2555183; PIDN:AAB81526.1

Query Match 4.4%; Score 7; DB 2; Length 1822;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 SSIDKQG 146  
|||||||  
Db 210 SSIDKQG 216

RESULT 30  
B49347  
DNA replication/recombination protein RecF - Caulobacter crescentus (fragment)  
C;Species: Caulobacter crescentus  
C;Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 03-May-1994  
C;Accession: B49347  
R;Rizzo, M.F.; Shapiro, L.; Gober, J.  
J. Bacteriol. 175, 6970-6981, 1993  
A;Title: Asymmetric expression of the gyrase B gene from the replication-competent ch  
A;Reference number: A49347; MUID:94042862  
A;Accession: B49347  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-56 <RIZ>  
A;Cross-references: GB:U00593

Query Match 3.8%; Score 6; DB 2; Length 56;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 GGRSVY 114  
|||||||  
Db 25 GGRSVY 30

RESULT 31  
D35054  
MHC Class II histocompatibility antigen HLA-DQ-1.4 beta chain - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 03-Aug-1990 #sequence\_revision 03-Aug-1990 #text\_change 23-Jul-1999  
C;Accession: D35054; I68814  
R;Gyllenstein, U.B.; Lashkari, D.; Erlich, H.A.  
Proc. Natl. Acad. Sci. U.S.A. 87, 1835-1839, 1990  
A;Title: Allelic diversification at the class II DQB locus of the mammalian major his  
A;Reference number: A35054; MUID:90175391  
A;Accession: D35054  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-57 <GYL>  
R;Bugawan, T.L.; Erlich, H.A.  
Immunogenetics 33, 163-170, 1991  
A;Title: Rapid typing of HLA-DQB1 DNA polymorphism using nonradioactive oligonucleoti  
A;Reference number: I54476; MUID:91184857  
A;Accession: I68814  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-57 <BUG>  
A;Cross-references: GB:M65047; NID:g187958; PIDN:AAA36246.1; PID:g187959  
C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 3.8%; Score 6; DB 2; Length 57;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 EEDVRF 89  
 Db 15 EEDVRF 20

## RESULT 32

S76286  
 hypothetical protein - *Synechocystis* sp. (strain PCC 6803)  
 C:Species: *Synechocystis* sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
 C:Accession: S76286  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
 S.  
 A:Reference number: S74322; MUID:97061201  
 A:Accession: S76286  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-66 <KAN>  
 A:Cross-references: EMBL:D64000; GB:AB001339; NID:gl001484; PIDN:BAA10138.1; PID:g167333  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 3.8%; Score 6; DB 2; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 RYPRYL 67  
 Db 10 RYPRYL 15

## RESULT 33

BNRT1  
 brain neuron cytoplasmic protein 1 - rat (fragment)  
 C:Species: *Rattus norvegicus* (Norway rat)  
 C:Date: 19-Feb-1984 #sequence\_revision 19-Feb-1984 #text\_change 22-Jun-1999  
 C:Accession: A03136  
 R:Sutcliffe, J.G.; Milner, R.J.; Shinnick, T.M.; Bloom, F.E.  
 Cell 33, 671-682, 1983  
 A:Title: Identifying the protein products of brain-specific genes with antibodies to che  
 A:Reference number: A90836; MUID:83259254  
 A:Accession: A03136  
 A:Molecule type: mRNA  
 A:Residues: 1-83 <SUT>  
 A:Cross-references: GB:V01543; GB:J00755; NID:g56876; PIDN:CAA24784.1; PID:g56877  
 A:Experimental source: clone pIA75  
 C:Superfamily: brain neuron cytoplasmic protein 1  
 C:keywords: brain; cytosol

Query Match 3.8%; Score 6; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 RSVSPW 52  
 Db 57 RSVSPW 62

## RESULT 34

T32439  
 hypothetical protein C30E1.1 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T32439  
 R:Blanchard, M.; Stellyes, L.; Beck, C.  
 submitted to the EMBL Data Library, September 1997  
 A:Description: The sequence of *C. elegans* cosmid C30E1.  
 A:Reference number: Z21167

A:Accession: T32439  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-85 <BLA>  
 A:Cross-references: EMBL:AF026204; PIDN:AAB71254.1; GSPDB:GN00028; CESP:C30E1.1  
 A:Experimental source: strain Bristol N2; clone C30E1  
 C:Genetics:  
 A:Gene: CESP:C30E1.1  
 A:Map position: X  
 A:Introns: 36/1

Query Match 3.8%; Score 6; DB 2; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 MPTVVL 101  
 Db 37 MPTVVL 42

## RESULT 35

T14922  
 hypothetical protein Y1005 - *Yersinia pestis* plasmid pMT1  
 C:Species: *Yersinia pestis*  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 22-Oct-1999  
 C:Accession: T14922  
 R:Lindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.  
 Infect. Immun. 66, 5731-5742, 1998  
 A:Title: Complete DNA sequence and detailed analysis of the *Yersinia pestis* KIM5 pla  
 A:Reference number: Z18268; MUID:99043898  
 A:Accession: T14922  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-90 <LIN>  
 A:Cross-references: EMBL:AF074611; NID:g3883003; PID:g3883007; PIDN:AAC82667.1  
 C:Genetics:  
 A:Gene: Y1005  
 A:Genome: plasmid pMT1

Query Match 3.8%; Score 6; DB 2; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 RGCLTG 80  
 Db 8 RGCLTG 13

## RESULT 36

JQ0527  
 capsid assembly protein 31-2 - phage T4  
 C:Species: phage T4  
 A:Note: host *Escherichia coli*  
 C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 08-Oct-1999  
 C:Accession: JQ0527; JQ0291; S26172  
 R:Pillipov, A.G.; Mesyanzhinov, V.V.; Aebi, U.; Kellenberger, E.  
 Nucleic Acids Res. 18, 3635, 1990  
 A:Title: Cloning and sequencing of bacteriophage T4 genes between map positions 128.  
 A:Reference number: JQ0524; MUID:90301484  
 A:Accession: JQ0527  
 A:Status: translation not shown

A:Molecule type: DNA  
 A:Residues: 1-97 <PRI>  
 A:Cross-references: EMBL:X17657; NID:gl5204; PIDN:CAA35653.1; PID:gl5209  
 R:Raudonikienė, A.; Nivinskis, R.  
 Nucleic Acids Res. 18, 4280, 1990  
 A:Title: Nucleotide sequence of bacteriophage T4 gene 31 region.  
 A:Reference number: JQ0290; MUID:90332452  
 A:Accession: JQ0291  
 A:Status: translation not shown  
 A:Molecule type: DNA

A;Residues: 26-97 <RAU>  
A;Cross-references: GB:M37882; GB:M23722; NID:g215873; PIDN:AAA32508.1; PTD:g215878  
R;Raudoniklene, A.; Nivinskis, R.  
Gene 114, 85-90, 1992  
A;Title: Gene rIII is the nearest downstream neighbour of bacteriophage T4 gene 31.  
A;Reference number: S26167; MUID:92267389  
A;Accession: S26172  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 26-97 <RA2>  
A;Cross-references: EMBL:X54536; NID:g15789; PIDN:CAA38407.1; PID:g15794

Query Match 3.8%; Score 6; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RARAVL 9  
| | | | |  
DB 80 RARAVL 85

RESULT 37  
T35604  
hypothetical protein SC6G9.08c - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
R;Accession: T35604  
R;Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, June 1999  
A;Reference number: Z21584  
A;Accession: T35604  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-106 <SEE>  
A;Cross-references: EMBL:AL079356; PIDN:CAB45599.1; GSPDB:GN00070; SCOEDB:SC6G9.08c  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCOEDB:SC6G9.08c

Query Match 3.8%; Score 6; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 98;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SARARA 7  
| | | | |  
DB 54 SARARA 59

RESULT 38  
T30659  
hypothetical protein 57L - Molluscum contagiosum virus 1  
N;Alternate names: MC057L  
C;Species: Molluscum contagiosum virus 1  
C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jul-2000  
R;Accession: T30659  
R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.  
Science 273, 813-816, 1996  
A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re  
A;Reference number: Z20876; MUID:96325459  
A;Accession: T30659  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-108 <SEN>  
A;Cross-references: EMBL:U60315; NID:g1491943; PIDN:RAC55185.1; PID:g1492000  
C;Genetics:  
A;Note: MC057L

Query Match 3.8%; Score 6; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ARARAV 8  
| | | | |  
DB 73 ARARAV 78

RESULT 39  
B81390  
probable membrane protein Cj0455c [imported] - Campylobacter jejuni (strain NCTC 1116  
C;Species: Campylobacter jejuni  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 28-Jul-2000  
R;Accession: B81390  
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chl  
N.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba  
Nature 403, 665-668, 2000  
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals  
A;Reference number: A81250; MUID:20150912  
A;Accession: B81390  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-114 <PAR>  
A;Cross-references: GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CAB75093.1; PID:g696  
A;Experimental source: serotype O2, strain NCTC 11168  
C;Genetics:  
C;Superfamily: Campylobacter jejuni probable membrane protein Cj0455c

Query Match 3.8%; Score 6; DB 2; Length 114;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 INSSID 143  
| | | | |  
DB 46 INSSID 51

RESULT 40  
T37050  
probable integral membrane protein - Streptomyces coelicolor (fragment)  
C;Species: Streptomyces coelicolor  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
R;Accession: T37050  
R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1999  
A;Reference number: Z21620  
A;Accession: T37050  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-114 <SEE>  
A;Cross-references: EMBL:AL109747; PIDN:CAB52347.1; GSPDB:GN00070; SCOEDB:SCJ21.01c  
A;Experimental source: strain A3(2)  
C;Genetics:  
A;Gene: SCOEDB:SCJ21.01c

Query Match 3.8%; Score 6; DB 2; Length 114;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 AYTIP 122  
| | | | |  
DB 40 AYTIP 45

RESULT 41  
A72474  
hypothetical protein APE2434 - Aeropyrum pernix (strain K1)  
C;Species: Aeropyrum pernix  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
R;Accession: A72474  
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum  
A:Reference number: A72450; MUID:99310339  
A:Accession: A72474  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-118 <RAW>  
A:Cross-references: DDBJ:AF000064; NID:g5105945; PIDN:BAA81449.1; PID:d1045235; PID:g510  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE2434  
C:Superfamily: Aeropyrum pernix hypothetical protein APE2434

Query Match 3.8%; Score 6; DB 2; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RAVLSA 11  
|||||  
Db 79 RAVLSA 84

RESULT 42  
T40303  
hypothetical protein SPC36.08c - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40303  
R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duesterh  
submitted to the EMBL Data Library, May 1998  
A:Reference number: Z21919  
A:Accession: T40303  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-120 <LYN>  
A:Cross-references: EMBL:AL023589; PIDN:CAA19056.1; GSPDB:GN00067; SPDB:SPBC36.08c  
A:Experimental source: strain 972h-; cosmid c36  
C:Genetics:  
A:Gene: SPDB:SPBC36.08c  
A:Map position: 2

Query Match 3.8%; Score 6; DB 2; Length 120;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 EKDADS 137.  
|||||  
Db 3 EKDADS 8

RESULT 43  
A24255  
chorion class A protein L1 precursor - silkworm  
C:Species: Bombyx mori (silkworm)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 20-Jun-2000  
C:Accession: A24255  
R:Spoerel, N.; Nguyen, H.T.; Kafatos, F.C.  
J. Mol. Biol. 190, 23-35, 1986  
A:Title: Gene regulation and evolution in the chorion locus of Bombyx mori. Structural a  
A:Reference number: A92929; MUID:87060979  
A:Accession: A24255  
A:Molecule type: DNA  
A:Residues: 1-129 <SPO>  
A:Cross-references: GB:X04028; GB:X04029; GB:X04030; GB:X04031; GB:X04032; GB  
C:Superfamily: chorion class A protein pc292  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-129/Product: chorion class A protein L1 #status predicted <MAT>

Query Match 3.8%; Score 6; DB 2; Length 129;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PACAGG 110  
|||||  
Db 103 PACAGG 108

RESULT 44  
T27248  
hypothetical protein Y5F2A.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T27248  
R:Lennard, N.  
submitted to the EMBL Data Library, October 1998  
A:Reference number: Z20331  
A:Accession: T27248  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-131 <WIL>  
A:Cross-references: EMBL:AL032641; PIDN:CAA21646.1; GSPDB:GN00022; CESP:Y5F2A.2  
A:Experimental source: clone Y5F2A  
C:Genetics:  
A:Gene: CESP:Y5F2A.2  
A:Map position: 4  
A:Introns: 50/1; 93/3  
C:Superfamily: Caenorhabditis hypothetical protein C40H1.5

Query Match 3.8%; Score 6; DB 2; Length 131;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LFGEED 86  
|||||  
Db 69 LFGEED 74

RESULT 45  
T09626  
outer membrane protein - Klebsiella pneumoniae (fragment)  
C:Species: Klebsiella pneumoniae  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T09626  
R:George, A.M.; Hall, R.M.; Stokes, H.W.  
Microbiology 141, 1909-1920, 1995  
A:Title: Multidrug resistance in Klebsiella pneumoniae: a novel gene, ramA, confers  
A:Reference number: Z16785; MUID:96032015  
A:Accession: T09626  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-132 <GEO>  
A:Cross-references: EMBL:U19581; NID:g885876; PIDN:AAA85696.1; PID:g885877  
A:Experimental source: strain ECL8Mdr  
C:Genetics:  
A:Gene: roma

Query Match 3.8%; Score 6; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RARAVL 9  
|||||  
Db 95 RARAVL 100

RESULT 46  
WMV265  
14K cell fusion protein - vaccinia virus (strain WR, 65-16)  
C:Species: vaccinia virus  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jul-1999  
C:Accession: A37076

R;Gong, S.; Lai, C.; Esteban, M.  
Virology 178, 81-91, 1990

A;Title: Vaccinia virus induces cell fusion at acid pH and this activity is mediated by  
A;Reference number: A37076; MUID:90357795

A;Accession: A37076

A;Molecule type: DNA

A;Residues: 1-136 <GON>

A;Cross-references: EMBL:M37086; NID:g335300; PIDN:AAA47961.1; PID:g335301

C;Superfamily: vaccinia virus 14K cell fusion protein

C;Keywords: glycoprotein; membrane fusion

F;2-136/Product: 14K cell fusion protein #status predicted <MAT>

F;86/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.8%; Score 6; DB 1; Length 136;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 PARYPR 65

Db 42 PARYPR 47

RESULT 47

T09583

acyl carrier protein - swamp oak

C;Species: Casuarina glauca (swamp oak)

C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C;Accession: T09583

R;Bogusz, D.

submitted to the EMBL Data Library, February 1997

A;Description: cDNA sequence for an acyl carrier protein from actinorhizal nodules of Ca

A;Reference number: Z16750

A;Accession: T09583

A;Status: preliminary; translated from GB/EMBL/DDBT

A;Molecule type: mRNA

A;Residues: 1-136 <BOG>

A;Cross-references: EMBL:Y10994

C;Superfamily: acyl carrier protein; acyl carrier protein homology

C;Keywords: carrier protein; fatty acid biosynthesis; phosphopantetheine; phosphoprotein

F;56-127/Domain: acyl carrier protein homology <ACP>

F;91/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 3.8%; Score 6; DB 2; Length 136;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 NLRSVS 50

Db 25 NLRSVS 30

RESULT 48

G75482

hypothetical protein - Deinococcus radiodurans (strain R1)

C;Species: Deinococcus radiodurans

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C;Accession: G75482

R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A;Reference number: A75250; MUID:20036896

A;Accession: G75482

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-136 <WHI>

A;Cross-references: GB:AE001929; GB:AE000513; NID:g6458437; PIDN:AAF10317.1; PID:g645844

A;Experimental source: strain R1

C;Genetics:

A;Gene: DR0736

A;Map position: 1

Query Match 3.8%; Score 6; DB 2; Length 136;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SARARA 7

Db 79 SARARA 84

RESULT 49

JH0127

mobA protein - Escherichia coli plasmid RSF1010

C;Species: Escherichia coli

C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 08-Oct-1999

C;Accession: JH0127; PS0290; S10917

R;Scholz, P.; Haring, V.; Wittmann-Liebold, B.; Ashman, K.; Bagdasarian, M.; Scherzin

Gene 75, 271-288, 1989

A;Title: Complete nucleotide sequence and gene organization of the broad-host-range p

A;Reference number: JH0123; MUID:89232758

A;Accession: JH0127

A;Molecule type: DNA

A;Residues: 1-137 <SCH>

A;Cross-references: GB:M28829; NID:gl52577; PIDN:AAA26446.1; PID:gl52582

A;Accession: PS0290

A;Molecule type: protein

A;Residues: 1-36 <SC2>

R;Derbyshire, K.M.; Hatfull, G.; Willetts, N.

Mol. Gen. Genet. 206, 161-168, 1987

A;Title: Mobilization of the non-conjugative plasmid RSF1010: A genetic and DNA seque

A;Reference number: S07319; MUID:87201082

A;Accession: S10917

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-137 <DER>

A;Cross-references: EMBL:X04830; NID:g42531; PIDN:CAA28521.1; PID:g42535

C;Comment: The plasmid RSF1010 is a naturally occurring broad-host-range plasmid belo

C;Comment: This protein is required for mobilization of plasmid RSF1010 in the presen

C;Genetics:

A;Gene: mobB

A;Genome: plasmid

Query Match 3.8%; Score 6; DB 2; Length 137;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 MPTVWL 101

Db 105 MPTVWL 110

RESULT 50

JC4628

cytotoxic T-lymphocyte-associated antigen 8 precursor - mouse

N;Alternate names: CTLA8 protein

C;Species: Mus musculus (house mouse)

C;Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 26-Aug-1999

C;Accession: JC4628

R;Yao, Z.; Timour, M.; Painter, S.; Fanslow, W.; Spriggs, M.

Gene 168, 223-225, 1996

A;Title: Complete nucleotide sequence of the mouse CTLA8 gene.

A;Reference number: JC4628; MUID:96194901

A;Accession: JC4628

A;Molecule type: DNA

A;Residues: 1-147 <YAO>

A;Cross-references: GB:U35108; NID:gl244499; PIDN:AAA93253.1; PID:gl244500

C;Genetics:

A;Gene: ctla8

A;Introns: 69/2

C;Superfamily: salmivi herpesvirus immediate-early protein 2

C;Keywords: cytokine; glycoprotein; lymphocyte  
F;1-14/Domain: signal sequence #status predicted <SIG>  
F;15-147/Product: cytotoxic T-lymphocyte-associated antigen 8 #status predicted <MAT>  
F;60/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.8%; Score 6; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 123 VGCTCV 128  
Db 134 VGCTCV 139

Search completed: April 14, 2002, 12:42:59  
Job time: 126 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2002, 12:42:33 ; Search time 11.85 Seconds  
(without alignments)  
495.053 Million cell updates/sec

Title: US-09-731-816-4

Perfect score: 160

Sequence: 1 NSARARAVLSAFHHTLQLGP.....SIDKQAKLLIGPNDAPAGP 160

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 6

Total number of hits satisfying chosen parameters: 178

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	5.0	837	1 YLE5_CAEEL	P46941 caenorhabdi
2	7	4.4	61	1 RS14_MYCPE	P47410 mycoplasma
3	7	4.4	61	1 RS14_MYCPN	Q50305 mycoplasma
4	7	4.4	275	1 PSB0_CYAA5	Q9R6W6 cyanothece
5	7	4.4	348	1 STR1_STRGR	P09400 streptomyce
6	7	4.4	360	1 HIS8_LACLA	Q02135 lactococcus
7	7	4.4	395	1 YQ37_CAEEL	Q09458 caenorhabdi
8	7	4.4	492	1 DYJ2_HUMAN	O43237 homo sapien
9	7	4.4	497	1 DYJ2_RAT	Q62698 rattus norv
10	7	4.4	582	1 MM14_RABIT	Q95220 oryctolagus
11	7	4.4	712	1 DNLJ_RHOMR	P49421 rhodothermu
12	7	4.4	972	1 UVRA_MYCTU	P94972 mycobacteri
13	6	3.8	83	1 NSG1_RAT	Q02683 rattus norv
14	6	3.8	97	1 Y13A_BPT4	P17310 bacterioph
15	6	3.8	129	1 CHAI_BOMMO	P08826 bombyx mori
16	6	3.8	132	1 ROMA_KLEPN	Q48412 klebsiella
17	6	3.8	136	1 VFUS_VACC6	P26312 vaccinia vi
18	6	3.8	137	1 MBB2_ECOLI	P07113 escherichia
19	6	3.8	150	1 IL17_RAT	Q61453 rattus norv
20	6	3.8	151	1 VG13_HSVSA	P24916 herpesvirus
21	6	3.8	151	1 IL17_HUMAN	Q16552 homo sapien
22	6	3.8	156	1 EST1_MOUSE	P11374 mus musculu
23	6	3.8	158	1 IL17_MOUSE	Q62386 mus musculu
24	6	3.8	174	1 SODC_CAUCR	P20379 caulobacter
25	6	3.8	177	1 Y13A_LACJO	Q48585 lactobacill
26	6	3.8	185	1 NSG1_HUMAN	P42857 homo sapien
27	6	3.8	185	1 NSG1_MOUSE	Q62092 mus musculu
28	6	3.8	186	1 MGH_BORBU	O51383 borrelia bu
29	6	3.8	199	1 UBC4_DROME	P52486 drosophila
30	6	3.8	207	1 HIS5_AQUAE	Q66943 aquifex ae
31	6	3.8	207	1 UCRI_CHRVI	O31214 chromatium
32	6	3.8	211	1 YOR3_SOUV3	Q04550 southampton
33	6	3.8	220	1 GT29_FASHE	P56598 fasciola he

34	3.8	232	1 C1B3_SHEEP	P80943 ovis aries
35	3.8	232	1 ELA_ADECT	P14264 canine aden
36	3.8	233	1 YSR3_CAEEL	Q09951 caenorhabdi
37	3.8	254	1 RLA_METH	O26111 methanobact
38	3.8	258	1 BPHF_RHOSO	O05151 rhodococcus
39	3.8	261	1 YHEB_CHLVI	P56160 chlorobium
40	3.8	262	1 OTCC_NEICU	Q01322 neisseria c
41	3.8	262	1 OTCC_NEIMU	Q01326 neisseria m
42	3.8	264	1 FTSQ_STRCO	P45518 streptomyce
43	3.8	271	1 YEDP_ECOLI	P76329 escherichia
44	3.8	291	1 LIP_THELA	O59952 thermomyces
45	3.8	294	1 YX95_MYCTU	O50730 mycobacteri
46	3.8	295	1 MPR1_SCHPO	O94321 schizosacch
47	3.8	302	1 YBHK_ECOLI	P75767 escherichia
48	3.8	311	1 SECF_RICPR	O92634 rickettsia
49	3.8	312	1 O2J2_HUMAN	O76002 homo sapien
50	3.8	317	1 CASB_HUMAN	O9Y240 homo sapien
51	3.8	317	1 CASB_MOUSE	O9Y240 mus musculu
52	3.8	321	1 BIRA_ECOLI	P06709 escherichia
53	3.8	322	1 YAUB_SCHPO	Q10166 schizosacch
54	3.8	326	1 RBSR_BACSU	P36944 bacillus su
55	3.8	328	1 Y350_MYCPN	P75252 mycoplasma
56	3.8	331	1 QOR_MOUSE	P47199 mus musculu
57	3.8	332	1 HEM2_PROFR	P77923 propionibac
58	3.8	336	1 VPB_BHP1	P51720 bacterioph
59	3.8	336	1 YDBI_SCHPO	Q10368 schizosacch
60	3.8	344	1 RUVB_MYCTU	O50629 mycobacteri
61	3.8	346	1 KMOS_CERAE	P10850 cercopithec
62	3.8	346	1 KMOS_HUMAN	P00340 homo sapien
63	3.8	348	1 CDNC_MOUSE	P49919 mus musculu
64	3.8	348	1 XKDT_BACSU	P54339 bacillus su
65	3.8	348	1 YQBT_BACSU	P45935 bacillus su
66	3.8	349	1 RUVB_MYCLE	O40833 mycobacteri
67	3.8	360	1 YAG3_MYCTU	O53411 mycobacteri
68	3.8	366	1 MTS2_METJA	Q58600 methanococc
69	3.8	381	1 CHYN_BOVIN	P00794 bos taurus
70	3.8	381	1 CPXG_STRSQ	P23296 streptomyce
71	3.8	382	1 KCC1_METAN	O14408 metahizium
72	3.8	387	1 RECF_CAUCR	P49998 caulobacter
73	3.8	396	1 NR21_ORYLA	O9Y913 oryzias lat
74	3.8	398	1 MPK2_CHICK	Q90891 gallus gall
75	3.8	403	1 GCST_CANFA	Q9T827 canis famli
76	3.8	403	1 GCST_HUMAN	P48728 homo sapien
77	3.8	404	1 PDUW_SALTY	P74879 salmonella
78	3.8	404	1 TRPB_THETH	P16609 thermus aqu
79	3.8	408	1 YOC2_YEAST	P25040 saccharomyc
80	3.8	409	1 PEXC_PICPA	Q01961 pichia past
81	3.8	410	1 PHT4_PSEPU	Q05184 pseudomonas
82	3.8	417	1 AVTA_ECOLI	P09053 escherichia
83	3.8	421	1 CGA1_MOUSE	O61456 mus musculu
84	3.8	427	1 IM44_SCHPO	O60084 schizosacch
85	3.8	427	1 NCAP_PIRYV	P26037 piry virus.
86	3.8	441	1 HISX_STRCO	P16245 streptomyce
87	3.8	447	1 SLF1_YEAST	Q12034 saccharomyc
88	3.8	452	1 PH4H_HUMAN	P00439 homo sapien
89	3.8	453	1 PH4H_MOUSE	P16331 mus musculu
90	3.8	453	1 PH4H_RAT	P04176 rattus norv
91	3.8	462	1 MYCN_MOUSE	P03966 mus musculu
92	3.8	474	1 SYFA_ARCFU	O28324 archaeoglob
93	3.8	477	1 SYC_CHLTR	O84787 chlamydia t
94	3.8	479	1 EFTI_SOYBN	Q43467 glycine max
95	3.8	490	1 CHEI_HUMAN	O00409 homo sapien
96	3.8	496	1 CAMA_HUMAN	P21941 homo sapien
97	3.8	509	1 PSBB_CVAPA	P20093 anabaena sp
98	3.8	509	1 PSBB_CVAPA	P48103 cyanophora
99	3.8	518	1 BAR2_SCHCO	Q05659 schizophyll
100	3.8	543	1 FLGR_BUCAI	P57428 buchnera ap
101	3.8	556	1 NUMB_DROME	P16554 drosophila
102	3.8	559	1 SUCP_ECOLI	P76041 escherichia
103	3.8	569	1 YSV1_YEAST	P24088 saccharomyc
104	3.8	581	1 YD15_SCHPO	Q10238 schizosacch
105	3.8	583	1 SHC_HUMAN	P29353 homo sapien
106	3.8	584	1 Y328_CHLPN	Q92810 chlamydia p

107	6	3.8	590	1	CHLL_ARATH	Q05085 arabidopsis	RESULT 1
108	6	3.8	610	1	TOH2_CAEEL	P98060 caenorhabdi	YLE5_CAEEL
109	6	3.8	616	1	SECD_HAEIN	P44591 haemophilus	ID YLE5_CAEEL
110	6	3.8	639	1	VGLG_SYN	P27277 sonchus yel	AC P46941
111	6	3.8	639	1	KPCI_DROME	P05130 drosophila	DT 01-NOV-1995 (Rel. 32, Created)
112	6	3.8	652	1	CN16_YEREN	P53052 yersinia en	DT 01-NOV-1995 (Rel. 32, Last sequence update)
113	6	3.8	676	1	YOTY_YEAST	Q08647 saccharomyc	DT 20-AUG-2001 (Rel. 40, Last annotation update)
114	6	3.8	688	1	MEAA_METEX	Q49115 methylobact	DE HYPOTHETICAL 94.2 KDA PROTEIN C38D4.5 IN CHROMOSOME III.
115	6	3.8	705	1	ICAL_BOVIN	P20811 bos taurus	GN C38D4.5.
116	6	3.8	716	1	PROL_ARATH	P43299 arabidopsis	OS Caenorhabditis elegans.
117	6	3.8	723	1	ICAL_SHEEP	Q95208 ovis aries	OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
118	6	3.8	728	1	BPHY_PSEAE	Q02987 rhodobacter	OC Rhabditidae; Peloderinae; Caenorhabditis.
119	6	3.8	739	1	HYPE_HROCA	P17201 acetobacter	OX NCBI_TaxID=6239;
120	6	3.8	751	1	PHYE_ANASP	P09248 varicella-z	[1]
121	6	3.8	752	1	HEPA_HSV2H	Q43950 azotobacter	SEQUENCE FROM N.A.
122	6	3.8	773	1	DHAQ_ACEPO	P40596 azotobacter	RC STRAIN-BRISTOL N2;
123	6	3.8	775	1	RIRL_VZVD	Q05036 caenorhabdi	RA Coles L.;
124	6	3.8	776	1	HYPE_AZCH	O29753 archaeoglob	RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
125	6	3.8	776	1	HYPE_AZOV1	P21580 homo sapien	CC - - SIMILARITY: C-TERMINAL TO CHIMAERIN.
126	6	3.8	776	1	YLA4_CAEEL	Q00839 homo sapien	CC - - SIMILARITY: CONTAINS 1 WW DOMAIN.
127	6	3.8	781	1	DPOL_ARCFU	P39001 saccharomyc	CC - - SIMILARITY: CONTAINS 1 WW DOMAIN.
128	6	3.8	781	1	TNP3_HUMAN	Q02410 homo sapien	CC This SWISS-PROT entry is copyright. It is produced through a collaboration
129	6	3.8	794	1	Z148_HUMAN	Q04671 homo sapien	CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
130	6	3.8	794	1	Z148_MOUSE	P33744 clostridium	CC the European Bioinformatics Institute. There are no restrictions on its
131	6	3.8	794	1	Z148_RAT	P07572 simian maso	CC use by non-profit institutions as long as its content is in no way
132	6	3.8	805	1	DF19_CAEEL	P04025 simian retr	CC modified and this statement is not removed. Usage by and for commercial
133	6	3.8	824	1	ROU_HUMAN	P51517 simian retr	CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
134	6	3.8	836	1	UME6_YEAST	Q14832 homo sapien	CC or send an email to license@isb-sib.ch).
135	6	3.8	837	1	APB1_HUMAN	O84882 chlamydia t	-----
136	6	3.8	838	1	P_HUMAN	P03160 woodchuck h	EMBL; Z46241; CAA86318.1; -
137	6	3.8	862	1	ADHE_CLOAB	P06958 escherichia	DR WormPep; C38D4.5; CE00918.
138	6	3.8	867	1	POL_MPMV	P57301 buchnera ap	DR InterPro; IPR001849; PH.
139	6	3.8	867	1	POL_SRV1	P03744 bacterioph	DR InterPro; IPR001198; RhoGAP.
140	6	3.8	867	1	POL_SRV2	P15938 saccharomyc	DR InterPro; IPR001202; WW.
141	6	3.8	877	1	MGR3_HUMAN	Q03749 bacillus th	DR Pfam; PF00169; PH; 1.
142	6	3.8	878	1	PMPI_CHLTR	Q45707 bacillus th	DR Pfam; PF00620; RhoGAP; 1.
143	6	3.8	879	1	DPOL_WHV1	Q45707 bacillus th	DR Pfam; PF00397; WW; 1.
144	6	3.8	886	1	ODP1_ECOLI	Q45708 bacillus th	DR SMART; SM00233; PH; 1.
145	6	3.8	887	1	ODP1_BUCAI	P18474 grapevine f	DR SMART; SM00324; RhoGAP; 1.
146	6	3.8	1026	1	VG37_BPT4	O13798 schizosacch	DR SMART; SM00456; WW; 1.
147	6	3.8	1071	1	PR16_YEAST	P42175 bacillus su	DR PROSITE; PS01159; WW_DOMAIN_1; 1.
148	6	3.8	1138	1	C7AA_BACTU	P11722 gallus gall	DR PROSITE; PS50020; WW_DOMAIN_2; 1.
149	6	3.8	1138	1	C7AB_BACUA	Q10451 schizosacch	DR PROSITE; PS50003; PH_DOMAIN; 1.
150	6	3.8	1138	1	C7AB_BACUK	P41004 schizosacch	DR PROSITE; PS50003; PH_DOMAIN; 1.
151	6	3.8	1184	1	POL2_GFLV	P28827 homo sapien	KW Hypothetical protein.
152	6	3.8	1202	1	YE01_SCHPO	P28828 mus musculu	FT DOMAIN 96 129 WW.
153	6	3.8	1228	1	NARG_BACSU	Q12879 homo sapien	FT DOMAIN 386 505 PH.
154	6	3.8	1256	1	FINC_CHICK	P35436 mus musculu	SQ SEQUENCE 837 AA; 94196 MW; D1895E622D1F5997 CRC64;
155	6	3.8	1261	1	YDEL_SCHPO	Q00959 rattus norv	
156	6	3.8	1324	1	CUT3_SCHPO	Q02004 dugbe virus	Query Match 5.0%; Score 8; DB 1; Length 837;
157	6	3.8	1452	1	PTPM_HUMAN	Q09779 schizosacch	Best Local Similarity 100.0%; Pred. No. 3.7;
158	6	3.8	1452	1	PTPM_MOUSE	P40105 saccharomyc	Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
159	6	3.8	1464	1	NME1_HUMAN	P53345 saccharomyc	
160	6	3.8	1464	1	NME1_MOUSE	P53819 saccharomyc	QY 43 PTNLSVS 50
161	6	3.8	1464	1	NME1_RAT	P51611 mesocricetu	
162	6	3.8	1551	1	VGLM_DUGBY	P13889 rubella vir	Db 312 PTNLSVS 319
163	6	3.8	1628	1	YABE_SCHPO	Q28343 canis fami	
164	6	3.8	1681	1	YE20_YEAST	P13887 ross river	RESULT 2
165	6	3.8	1859	1	YG67_YEAST	Q91740 xenopus lae	RS14_MYCGE
166	6	3.8	1859	1	YN79_YEAST	P39846 bacillus su	ID RS14_MYCGE
167	6	3.8	2090	1	HFCL_MESAU	P51592 drosophila	AC P47410;
168	6	3.8	2205	1	POLN_RUBVT	P03186 epstein-bar	DT 01-FEB-1996 (Rel. 33, Created)
169	6	3.8	2333	1	PGCA_CANFA	Q03132 saccharopol	DT 01-FEB-1996 (Rel. 33, Last sequence update)
170	6	3.8	2479	1	POLN_RRVN	P27206 bacillus su	DT 20-AUG-2001 (Rel. 40, Last annotation update)
171	6	3.8	2481	1	FINC_XENLA	Q04747 bacillus su	DE 30S RIBOSOMAL PROTEIN S14.
172	6	3.8	2560	1	PPS2_BACSU	Q05470 bacillus su	GN RPSN OR RPS14 OR MG164.
173	6	3.8	2895	1	HYDE_DROME		OS Mycoplasma genitalium.
174	6	3.8	3149	1	TEGU_EBV		
175	6	3.8	3567	1	ERY2_SACER		
176	6	3.8	3587	1	SRF1_BACSU		
177	6	3.8	3587	1	SRF2_BACSU		
178	6	3.8	4427	1	PKSL_BACSU		

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Karlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium.";  
 RL Science 270:397-403(1995).  
 CC -!- FUNCTION: KNOWN TO BE REQUIRED FOR THE ASSEMBLY OF 30S PARTICLES  
 CC AND MAY ALSO BE RESPONSIBLE FOR DETERMINING THE CONFORMATION OF  
 CC THE 16S RNA AT THE A SITE (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U39696; AAC71382.1; -;  
 DR TIGR; MG164; -;  
 DR InterPro: IPR001209; Ribosomal\_S14.  
 DR Pfam: PF00253; Ribosomal\_S14; 1.  
 DR PROSITE: PS00527; RIBOSOMAL\_S14; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 61 AA; 6517 MW; 0B04CA96D071CD8F CRC64;

Query Match 4.4%; Score 7; DB 1; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 RARAVLS 10  
 DB 29 RARAVLS 35

RESULT 3  
 RS14\_MYCPN STANDARD; PRT; 61 AA.  
 AC Q50305;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE 30S RIBOSOMAL PROTEIN S14.  
 GN RPSN OR MPN178 OR MP653.  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RX MEDLINE=96177562; PubMed=8604303;  
 RA Hilbert H., Himmelfreich R., Plagens H., Herrmann R.;  
 RT "Sequence analysis of 56 kb from the genome of the bacterium  
 RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and  
 RT a cluster of ribosomal protein genes.";  
 RL Nucleic Acids Res. 24:628-639(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,

RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -!- FUNCTION: KNOWN TO BE REQUIRED FOR THE ASSEMBLY OF 30S PARTICLES  
 CC AND MAY ALSO BE RESPONSIBLE FOR DETERMINING THE CONFORMATION OF  
 CC THE 16S RNA AT THE A SITE (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U34795; AAC43703.1; -;  
 DR EMBL; AE000061; AAB96301.1; -;  
 DR InterPro: IPR001209; Ribosomal\_S14.  
 DR Pfam: PF00253; Ribosomal\_S14; 1.  
 DR PROSITE: PS00527; RIBOSOMAL\_S14; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 61 AA; 6885 MW; 552BA0FF662C481D CRC64;

Query Match 4.4%; Score 7; DB 1; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 RARAVLS 10  
 DB 29 RARAVLS 35

RESULT 4  
 PSBO\_CYAA5 STANDARD; PRT; 275 AA.  
 AC Q9R6W6;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PHOTOSYSTEM II MANGANESE-STABILIZING POLYPEPTIDE PRECURSOR (MSP).  
 GN PSBO.  
 OS Cyanothece (strain ATCC 51142).  
 OC Bacteria; Cyanobacteria; Chroococcales; Cyanothece.  
 OX NCBI\_TaxID=43989;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Tucker D.L., Hirsh K.R., Sherman L.A.;  
 RT "Regulation of PSII extrinsic proteins and O2-evolution in the  
 RT unicellular, diazotrophic cyanobacterium Cyanothece sp. ATCC 51142.";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: MSP BINDS TO A PUTATIVE MN-BINDING PROTEIN AND KEEPS  
 CC 2 OF THE 4 MN-ATOMS ASSOCIATED WITH PSII (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE PSBO FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; AF201467; AAF13997.1; ALT\_INIT.  
 DR InterPro: IPR002628; MSP.  
 DR Pfam: PF01716; MSP; 2.  
 KW Photosynthesis; Photosystem II; Manganese; Signal.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT CHAIN 29 275 PHOTOSYSTEM II MANGANESE-STABILIZING  
 FT POLYPEPTIDE.  
 SQ SEQUENCE 275 AA; 29970 MW; 89912B01911413DB CRC64;

```
Query Match          4.4%; Score 7; DB 1; Length 275;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 SINSSID 143
DB 164 SINSSID 170
|||||

RESULT 5
STRI_STRGR          STANDARD;          PRT; 348 AA.
AC P09400;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE STREPTOMYCIN BIOSYNTHESIS PROTEIN STRI.
GN STRI.
OS Streptomyces griseus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae.
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2-3-11;
RA Mansouri K., Piepersberg W.;
RT "Genetics of streptomycin production in Streptomyces griseus:
RT nucleotide sequence of five genes, strFGHIK, including a phosphatase
RT gene.";
RL Mol. Gen. Genet. 228:459-469(1991).
CC 1- PATHWAY: STREPTOMYCIN BIOSYNTHESIS.
CC 1- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y00459; CAA68521.1; -
DR PIR; S17779; S17779.
DR InterPro; IPRO00683; GFO_IDH_MOCA.
DR Pfam; PF01408; GFO_IDH_MOCA; 1.
KW Streptomycin biosynthesis; NAD.
FT NP_BIND 2 32 NAD (BY SIMILARITY).
SQ SEQUENCE 348 AA; 36709 MW; B95BCC443EEABDF6 CRC64;

Query Match          4.4%; Score 7; DB 1; Length 348;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ARAVLSA 11
DB 217 ARAVLSA 223
|||||

RESULT 6
HIS8_LACLA          STANDARD;          PRT; 360 AA.
ID HIS8_LACLA
AC Q02135; Q34130;
DT 01-JUL-1993 (Rel. 26, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC 2.6.1.9) (IMIDAZOLE ACETOL-
DE PHOSPHATE TRANSAMINASE).
GN HIS8.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
```

```
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCDO 2118;
RA Delorme C., Ehrlich S.D., Renault P.;
RT "Histidine biosynthesis genes in Lactococcus lactis subsp. lactis.";
RL J. Bacteriol. 174:6571-6579(1992).
RN [2]
RP REVISIONS.
RA Delorme C., Goupil-Feuillerat N., Godon J.-J., Ehrlich S.D.,
RA Renault P.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBDJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC 1- CATALYTIC ACTIVITY: L-HISTIDINOL-PHOSPHATE + 2-OXOGLUTARATE -
CC 3-(IMIDAZOL-4-YL)-2-OXOPROPYL PHOSPHATE + GLUTAMATE.
CC 1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC 1- PATHWAY: EIGHTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC 1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U92974; AAB81901.1; -
DR PIR; AE006353; AAK05304.1; -.
DR InterPro; IPRO01917; Aminotransf_2.
DR Pfam; PF00222; aminotran_2; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
KW Histidine biosynthesis; Transferase; Aminotransferase;
KW Pyridoxal phosphate; Complete proteome.
FT BINDING 208 208 PYRIDOXAL PHOSPHATE (PROBABLE).
FT CONFLICT 24 24 I -> M (IN REF. 2).
FT CONFLICT 69 69 D -> E (IN REF. 2).
FT CONFLICT 79 79 I -> F (IN REF. 2).
FT CONFLICT 135 135 L -> S (IN REF. 2).
FT CONFLICT 152 152 D -> N (IN REF. 2).
FT CONFLICT 179 179 N -> D (IN REF. 2).
FT CONFLICT 304 304 E -> D (IN REF. 2).
SQ SEQUENCE 360 AA; 41400 MW; C747B78C3495AC1 CRC64;

Query Match          4.4%; Score 7; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 NLRVSP 51
DB 6 NLRVSP 12
|||||

RESULT 7
YQ37_CAEEL          STANDARD;          PRT; 395 AA.
ID YQ37_CAEEL
AC Q09458;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
```

DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL 45.3 KDA PROTEIN C09F5.7 IN CHROMOSOME II.  
GN C09G5.7.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
ON NCBI\_TaxID=6239;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Palmer S.;  
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: TO C-ELEGANS C54G4.2 AND F44D12.8.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z46791; CAA86759.1; -  
DR WormPep; C09G5.7; CE01487.  
DR Hypothetical protein.  
KW  
SQ SEQUENCE 395 AA; 45309 MW; A86A8F5540B3A64F CRC64;  
  
Query Match 4.4%; Score 7; DB 1; Length 395;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 53 AYRISYD 59  
DB 351 AYRISYD 357  
|||||||  
  
RESULT 8  
DYJ2\_HUMAN STANDARD; PRT; 492 AA.  
ID DYJ2\_HUMAN  
AC Q43237;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DYNEIN LIGHT INTERMEDIATE CHAIN 2, CYTOSOLIC (LIC53/55) (LIC-2).  
GN DNCL12 OR LIC2  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RX [1]  
RP SEQUENCE FROM N.A.  
RA Zha D., Hu G.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: MAY PLAY A ROLE IN REGULATING INTERACTIONS BETWEEN  
CC DYNEIN AND P150-GLUED, AND THE CELLULAR SUBSTRATES FOR DYNEIN-  
CC MEDIATED MOTILITY (SUCH AS ORGANELLES) (BY SIMILARITY).  
CC -!- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND  
CC GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF035812; AAB88513.1; -  
DR Motor protein; Microtubules; Dynein; ATP-binding.  
KW NP\_BIND 61 68  
FT ATP (POTENTIAL).  
DE (MEMBRANE-TYPE-1 MATRIX METALLOPROTEINASE 1) (MT-MMP1).  
SQ SEQUENCE 492 AA; 54099 MW; AF7B4E49E3983DCC CRC64;

Query Match 4.4%; Score 7; DB 1; Length 492;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 148 KLLGPN 154  
DB 9 KLLGPN 15  
|||||||  
  
RESULT 9  
DYJ2\_RAT STANDARD; PRT; 497 AA.  
ID DYJ2\_RAT  
AC Q62698;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DYNEIN LIGHT INTERMEDIATE CHAIN 2, CYTOSOLIC (LIC53/55) (LIC-2).  
GN DNCL12.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RX [1]  
RP SEQUENCE FROM N.A.  
RA Tissue=Brain;  
RC MEDLINE=95256330; PubMed=7738094;  
RA Hughes S.M., Vaughan K.T., Herskovits J.S., Vallee R.B.;  
RT "Molecular analysis of a cytoplasmic dynein light intermediate chain  
RT reveals homology to a family of ATPases."  
RL J. Cell Sci. 108:17-24(1995).  
CC -!- FUNCTION: MAY PLAY A ROLE IN REGULATING INTERACTIONS BETWEEN  
CC DYNEIN AND P150-GLUED, AND THE CELLULAR SUBSTRATES FOR DYNEIN-  
CC MEDIATED MOTILITY (SUCH AS ORGANELLES).  
CC -!- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND  
CC GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.  
CC -!- TISSUE SPECIFICITY: UBIQUITOUS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U15139; AAA80334.1; -  
DR Motor protein; Microtubules; Dynein; ATP-binding.  
KW NP\_BIND 61 68  
FT ATP (POTENTIAL).  
DE (MEMBRANE-TYPE-1 MATRIX METALLOPROTEINASE 1) (MT-MMP1).  
SQ SEQUENCE 497 AA; 54744 MW; D4E2715880E194B4 CRC64;  
  
Query Match 4.4%; Score 7; DB 1; Length 497;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 148 KLLGPN 154  
DB 9 KLLGPN 15  
|||||||  
  
RESULT 10  
MM14\_RABIT STANDARD; PRT; 582 AA.  
ID MM14\_RABIT  
AC Q95220; P79225;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MATRIX METALLOPROTEINASE-14 PRECURSOR (EC 3.4.24.-) (MMP-14)  
DE (MEMBRANE-TYPE-1 MATRIX METALLOPROTEINASE 1) (MT-MMP1) (MTMMP1)  
DE MM14.  
ON Oryctolagus cuniculus (Rabbit).  
OS

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
RN NCBI\_TaxID=9986;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Vascular smooth muscle;  
RA Wang H., Kelsier J.;  
RL Submitted (Feb-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 1-572 FROM N.A.  
RC STRAIN=NEW ZEALAND WHITE;  
RA Sato T.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: SEEMS TO SPECIFICALLY ACTIVATE OF PRO-GELATINASE A. MAY  
CC THIS TRIGGER INVASION BY TUMOR CELLS BY ACTIVATING PRO-GELATINASE  
CC A ON THE TUMOR CELL SURFACE (BY SIMILARITY).  
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U83918; AAB41500.1; -;  
DR EMBL; U73940; AAD13803.1; -;  
DR HSSP; P08254; LUMS.  
DR MEROPS; M10.014; -;  
DR InterPro; IPR000585; Hemopexin.  
DR InterPro; IPR001818; Matrixin.  
DR InterPro; IPR000130; Zn\_MTPetdse.  
DR Pfam; PF00045; hemopexin; 4.  
DR PRINTS; PR00138; Peptidase\_M10; 1.  
DR SMART; SM00120; HX; 4.  
DR SMART; SM00235; ZmMc; 1.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
KW Hydrolyase; Metalloprotease; zinc; zymogen; Calcium; Signal;  
KW Transmembrane.  
FT SIGNAL 1 20 POTENTIAL.  
FT PROPEP 21 111 ACTIVATION PEPTIDE.  
FT CHAIN 112 582 MATRIX METALLOPROTEINASE-14.  
FT DOMAIN 112 541 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 542 562 POTENTIAL.  
FT DOMAIN 563 582 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 316 511 HEMOPEXIN-LIKE.  
FT SITE 93 93 CYSTEINE SWITCH (POTENTIAL).  
FT METAL 239 239 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 240 240 ZINC (CATALYTIC) (BY SIMILARITY).  
FT METAL 243 243 ZINC (CATALYTIC) (BY SIMILARITY).  
FT METAL 249 249 ZINC (CATALYTIC) (BY SIMILARITY).  
FT DISULFID 319 508 BY SIMILARITY.  
FT CONFLICT 29 29 Q -> K (IN REF. 2).  
FT CONFLICT 268 268 K -> R (IN REF. 2).  
FT CONFLICT 270 270 L -> V (IN REF. 2).  
FT CONFLICT 275 275 E -> D (IN REF. 2).  
FT CONFLICT 292 296 RGLLN -> KMPPP (IN REF. 2).  
FT CONFLICT 298 300 GOP -> RTT (IN REF. 2).  
FT CONFLICT 302 308 GLLFRIS -> RTFFPK (IN REF. 2).  
FT CONFLICT 310 310 G -> R (IN REF. 2).  
FT CONFLICT 317 317 K -> N (IN REF. 2).  
FT CONFLICT 329 329 F -> L (IN REF. 2).  
FT CONFLICT 360 360 L -> F (IN REF. 2).  
SQ SEQUENCE 582 AA; 65963 MW; 844624B0AFLB6812 CRC64;

Query Match 4.4%; Score 7; DB 1; Length 582;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 CPAGGRP 35  
| | | | | | | |  
DB 508 CPAGGRP 514  
RESULT 11  
DNLJ\_RHOMR  
ID DNLJ\_RHOMR STANDARD; PRT; 712 AA.  
AC P49421;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DNA LIGASE (EC 6.5.1.2) (POLYDEOXYRIBONUCLEOTIDE SYNTHASE [NAD+]).  
OS LIGA OR LIG.  
GN Rhodothermus marinus (Rhodothermus obamensis).  
OC Bacteria; CFB group; Rhodothermus group; Rhodothermus.  
OX NCBI\_TaxID=29549;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=R-21;  
RA Thorbjarnardottir S.H., Jonsson Z.O., Andersson O.S.,  
RA Kristjansson J.K., Egertsson G., Palsdottir A.;  
RT "Cloning and sequence analysis of the DNA ligase-encoding gene of  
RT Rhodothermus marinus, and overproduction, purification and  
RT characterization of two thermophilic DNA ligases.";  
RL Gene 161:1-6(1995).  
CC -1- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER  
CC LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-  
CC STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR  
CC THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF  
CC DAMAGED DNA. THIS ENZYME IS THERMOSTABLE BEING ACTIVE AT 5-75  
CC DEGREES CELSIUS WITH APPARENT OPTIMAL ACTIVITY ABOVE 55 DEGREES  
CC CELSIUS.  
CC -1- CATALYTIC ACTIVITY: NAD(+) + (DEOXYRIBONUCLEOTIDE)(N) +  
CC (DEOXYRIBONUCLEOTIDE)(M) -> AMP + NICOTINAMIDE NUCLEOTIDE +  
CC (DEOXYRIBONUCLEOTIDE)(N+M).  
CC -1- MISCELLANEOUS: THIS ENZYME IS THERMOSTABLE.  
CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U10483; AAA93198.1; -;  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001679; DNA\_ligase\_N.  
DR InterPro; IPR003265; Endo\_3c.  
DR InterPro; IPR000445; HHH.  
DR InterPro; IPR003583; HHH\_1.  
DR Pfam; PF00533; BRCT; 1.  
DR Pfam; PF01653; DNA\_ligase\_N; 1.  
DR Pfam; PF00633; HHH; 1.  
DR ProDom; PD003944; DNA\_ligase\_N; 1.  
DR SMART; SM00292; BRCT; 1.  
DR SMART; SM00478; ENDO3c; 1.  
DR SMART; SM00278; HHH1; 3.  
DR SMART; SM00532; LIGANC; 1.  
DR PROSITE; PS50172; BRCT; 1.  
DR PROSITE; PS01055; DNA\_LIGASE\_N1; 1.  
DR PROSITE; PS01056; DNA\_LIGASE\_N2; 1.  
KW Ligase; DNA repair; DNA replication; NAD.  
FT BINDING 153 153 AMP (BY SIMILARITY).  
SQ SEQUENCE 712 AA; 79487 MW; 893AD3A78F77FEC1 CRC64;

Query Match 4.4%; Score 7; DB 1; Length 712;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 FGEDVR 88  
DB 120 FGEDVR 126

RESULT 12  
UVRA\_MYCTU STANDARD; PRT; 972 AA.  
ID UVRA\_MYCTU  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE EXCINUCLEASE ABC SUBUNIT A.  
GN UVRA OR RV1638 OR MT1675 OR MTCY06H11.02.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
[1]  
SEQUENCE FROM N.A.  
RP STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."  
RL Nature 393:537-544(1998).  
[2]  
SEQUENCE FROM N.A.  
RP STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).  
CC -!- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRB AND UVRC.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). CONTAINS TWO ABC DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; Z85982; CAB06633.1; -  
CC DR EMBL; AE007031; AAK45944.1; ALT\_INIT.  
CC DR TIGR; MT1675; -  
CC DR Tuberculist; RV1638; -  
CC DR InterPro; IPR003439; ABC\_transportr.

Pfam; PF00005; ABC\_tran; 2.  
KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;  
KW DNA-binding; Zinc-finger; Complete proteome.  
FT NP\_BIND 32 39  
ATP (POTENTIAL).  
FT NP\_BIND 654 661  
ATP (POTENTIAL).  
FT ZN\_FING 257 285  
C4-TYPE (ATYPICAL).  
FT ZN\_FING 753 779  
C4-TYPE.  
SQ SEQUENCE 972 AA; 106131 MW; 8937A764E592D981 CRC64;

Query Match 4.4%; Score 7; DB 1; Length 972;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 LGPREQA 24  
DB 462 LGPREQA 468

RESULT 13  
NSG1\_RAT STANDARD; PRT; 83 AA.  
ID NSG1\_RAT  
AC P02683;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE NEURON SPECIFIC PROTEIN FAMILY MEMBER 1 (BRAIN NEURON CYTOPLASMIC PROTEIN 1) (FRAGMENT).  
GN NSG1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
SEQUENCE FROM N.A.  
RP TISSUE=Brain;  
RX MEDLINE=83259254; PubMed=6347394;  
RA Sutcliffe J.G., Milner R.J., Shinnick T.M., Bloom F.E.;  
RT "Identifying the protein products of brain-specific genes with antibodies to chemically synthesized peptides."  
RL Cell 33:671-682(1983).  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE NSG FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; V01543; CAA24784.1; -  
CC DR PIR; A03136; BNRT1.  
KW Golgi stack; Neurone.  
FT NON\_TER 1  
FT DOMAIN <1 83  
LUMENAL (POTENTIAL).  
SQ SEQUENCE 83 AA; 9582 MW; 71324757B9F5E99D CRC64;

Query Match 3.8%; Score 6; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 RSVSPW 52  
DB 57 RSVSPW 62

RESULT 14  
Y13A\_BPT4 STANDARD; PRT; 97 AA.  
ID Y13A\_BPT4

```

AC P17310;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE HYPOPHYSICAL 11.1 KDA PROTEIN IN GP30-RIII INTERGENIC REGION (ORF D).
GN V13A OR 30.9 OR 31.-2.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D;
RX MEDLINE=90301484; PubMed=2362813;
RA Prilipov A.G., Mesyanzhinov V.V., Aebi U., Kellenberger E.;
RT "Cloning and sequencing of bacteriophage T4 genes between map
RT positions 128.3-130.3.";
RL Nucleic Acids Res. 18:3635-3635(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90332452; PubMed=2377483;
RA Raudonikienė A., Nivinskas R.;
RT "Nucleotide sequence of bacteriophage T4 gene 31 region.";
RL Nucleic Acids Res. 18:4280-4280(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92267389; PubMed=1587487;
RA Raudonikienė A., Nivinskas R.;
RT "Gene rIII is the nearest downstream neighbour of bacteriophage T4
RT gene 31.";
RL gene 114:85-90(1992).
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-26 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17657; CAA35653.1; -
DR EMBL; M37882; AAA32508.1; ALT_INIT.
DR EMBL; X34536; CAA38407.1; ALT_INIT.
DR PIR; JQ0527; JQ0527.
DR PIR; S26172; S26172.
KW Hypothetical protein.
SQ SEQUENCE 97 AA; 11087 MW; 50C87172C3C2CA51 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RARAVL 9
DB 80 RARAVL 85

RESULT 15
CHAL_BOMMO STANDARD; PRT; 129 AA.
AC F08826;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DE CHORION CLASS A PROTEIN L11 PRECURSOR.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=703;
RX MEDLINE=87060979; PubMed=3023635;
RA Spoerel N., Nguyen H.T., Kafatos F.C.;
RT "Gene regulation and evolution in the chorion locus of Bombyx mori..
RT Structural and developmental characterization of four eggshell genes
RT and their flanking DNA regions.";
RL J. Mol. Biol. 190:23-35(1986).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE
CC SILK MOTH.
CC -1- SIMILARITY: MEMBER OF THE ALPHA-BRANCH OF CHORION PROTEIN TO WHICH
CC BELONG CLASSES A, CA AND HCA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15557; CAA33565.1; -
DR PIR; A24255; A24255.
DR HSSP; P10968; 2CWG.
KW Eggshell; Chorion; Repeat; Multigene family; Signal.
FT SIGNAL 1 21
FT CHAIN 22 129 CHORION CLASS A PROTEIN L11.
FT DOMAIN 22 64 LEFT ARM.
FT DOMAIN 65 112 CENTRAL DOMAIN.
FT DOMAIN 113 129 RIGHT ARM.
SQ SEQUENCE 129 AA; 12156 MW; 7204F6EEC05ABEAA CRC64;

Query Match 3.8%; Score 6; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 PACAGG 110
DB 103 PACAGG 108

RESULT 16
ROMA_KLEPN STANDARD; PRT; 132 AA.
AC Q48412;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE OUTER MEMBRANE PROTEIN ROMA (FRAGMENT).
GN ROMA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECL8;
RX MEDLINE=96032015; PubMed=7551053;
RA George A.M., Hall R.M., Stokes H.W.;
RT "Multidrug resistance in Klebsiella pneumoniae: a novel gene, ramA,
RT confers a multidrug resistance phenotype in Escherichia coli.";
RL Microbiology 141:1909-1920(1995).
CC -1- SIMILARITY: TO M.TUBERCULOSIS RV0906.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; U19581; AAA85696.1; --
KW Outer membrane.
FT NON_TER 1
SQ SEQUENCE 132 AA; 14745 MW; BCB216F27F9C853B CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RARAVL 9
DB 95 RARAVL 100

RESULT 17
VFUS_VACC6
ID VFUS_VACC6 STANDARD; PRT; 136 AA.
AC P26312;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE 14 KDA FUSION PROTEIN.
GN A27L.
OS Vaccinia virus (strain WR 65-16).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10247;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90357795; PubMed=2389560;
RA Gong S., Lai C., Esteban M.;
RT "Vaccinia virus induces cell fusion at acid pH and this activity is
RL mediated by the N-terminus of the 14-kDa virus envelope protein.";
RL Virology 178:81-91(1990).
CC -!- FUNCTION: THIS PROTEIN APPEARS TO PLAY AN IMPORTANT ROLE IN VIRUS
CC PENETRATION AT THE LEVEL OF CELL FUSION. THE N-TERMINAL PROXIMAL
CC REGION IS ESSENTIAL FOR FUSION ABILITY. ESSENTIAL IN FUSING THE
CC OUTERMOST OF THE TWO GOLGI-DERIVED MEMBRANES ENVELOPING THE VIRUS
CC WITH THE PLASMA MEMBRANE, AND IN ITS SUBSEQUENT RELEASE
CC EXTRACELLULARLY.
CC -!- SUBUNIT: HOMOPRIMER, COVALENTLY LINKED.
CC -!- SUBCELLULAR LOCATION: ENVELOPE FRACTION OF VIRIONS.
CC -!- SIMILARITY: TO OTHER POXVIRUSES FUSION PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37086; AAA47961.1; --
DR PIR; A37076; MWV265.
DR InterPro; IPR003436; Vac_Fusion.
DR Pfam; PF02346; Vac_Fusion; 1.
KW Fusion protein; Glycoprotein.
FT CARBOHYD 86
SQ SEQUENCE 136 AA; 15837 MW; 5F685204E808D108 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 PARYPR 65
DB 42 PARYPR 47

RESULT 18
MBB2_ECOLI

```

```

ID MBB2_ECOLI STANDARD; PRT; 137 AA.
AC P07113;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MOBILIZATION PROTEIN B.
GN MOBB.
OS Escherichia coli.
OG Plasmid IncQ RSF1010.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87201082; PubMed=3033438;
RA Derbyshire K.M., Hatfull G., Willetts N.;
RT "Mobilization of the non-conjugative plasmid RSF1010: a genetic and
RL DNA sequence analysis of the mobilization region.";
RL Mol. Gen. Genet. 206:161-168(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232758; PubMed=2653965;
RA Scholz P., Haring V., Wittmann-Liebold B., Ashman K.,
RT "Complete nucleotide sequence and gene organization of the
RL broad-host-range plasmid RSF1010.";
RL Gene 75:271-288(1989).
CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO PROMOTE THE SPECIFIC
CC TRANSFER OF THE PLASMID IN THE PRESENCE OF CONJUGATIVE PLASMIDS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04830; CAA28521.1; --
DR EMBL; M28829; AAA26446.1; --
DR PIR; JH0127; JH0127.
DR PIR; SI0917; SI0917.
KW Mobility protein; Plasmid; Conjugation.
SQ SEQUENCE 137 AA; 15112 MW; D6ED69E1C3BB6910 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 MPTVVL 101
DB 105 MPTVVL 110

RESULT 19
IL17_RAT
ID IL17_RAT STANDARD; PRT; 150 AA.
AC Q61453;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED
DE ANTIGEN 8) (CTLA-8).
GN IL17 OR CTLA8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93294300; PubMed=8390535;
RA Rouvler E., Luciani M.F., Mattel M.-G., Denizot F., Golstein P.;

```

```
RT "CTLA-8, cloned from an activated T cell, bearing AU-rich messenger
RT RNA instability sequences, and homologous to a herpesvirus saimiri
RT gene."
RL J. Immunol. 150:5445-5456(1993).
RN [2]
RP ORGANISM IDENTIFICATION.
RX MEDLINE=96194901; PubMed=8654948;
RA Yao Z., Timour M., Painter S., Fanslow W., Spriggs M.K.;
RT "Complete nucleotide sequence of the mouse CTLA8 gene."
RL Gene 168:223-225(1996).
RN [3]
RP SEQUENCE FROM N.A., AND ORGANISM IDENTIFICATION.
RX MEDLINE=97031826; PubMed=887732;
RA Kennedy J., Rossi D.L., Zurawski S.M., Vega F. Jr., Kastelein R.A.,
RT Wagner J.L., Hannum C.H., Zlotnik A.;
RT CD4-CD8-T cells."
RL J. Interferon Cytokine Res. 16:611-617(1996).
CC -1- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND
CC HEMATOPOIETIC CYTOKINES (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: STRONG. TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO
CC HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE FROM MOUSE BUT, ON
CC THE BASIS OF SUBSEQUENT WORK (REF.2 AND REF.3), HAS BEEN SHOWN TO
CC BE OF RAT ORIGIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
RT binding protein.";
RL Virology 174:533-542(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333688; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021021; PubMed=1699352;
RA Nicholas J., Smith E.P., Coles L., Honess R.;
RT "Gene expression in cells infected with gammaherpesvirus saimiri:
RT properties of transcripts from two immediate-early genes.";
RL Virology 179:189-200(1990).
CC -1- SIMILARITY: STRONG. TO MAMMALIAN INTERLEUKIN-17 (CTLA-8).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

J. Exp. Med. 183:2593-2603(1996).  
[2]  
SEQUENCE FROM N.A.  
TISSUE-T-cell.  
MEDLINE-96094436; PubMed-7499828;  
Xao Z., Painter S.L., Fanslow W.C., Ulrich D., Macduff B.M.,  
Spriggs M.K., Armitage R.J.;  
"Human IL-17: a novel cytokine derived from T cells.";  
J. Immunol. 155:5483-5486(1995).  
CC -!- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND  
CC HEMATOPOIETIC CYTOKINES. ENHANCES THE SURFACE EXPRESSION OF THE  
CC INTRACELLULAR ADHESION MOLECULE-1 (ICAM-1) IN FIBROBLASTS.  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- TISSUE SPECIFICITY: RESTRICTED TO ACTIVATED MEMORY T-CELLS.  
CC -!- PTM: FOUND BOTH IN GLYCOSYLATED AND NONGLYCOSYLATED FORMS.  
CC -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO  
CC HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z58820; CAA91233.1; -  
DR EMBL; U32659; AAC50341.1; -  
DR MIM; 603149; -  
KW Cytokine; Glycoprotein; T-cell; Antigen; Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 155 INTERLEUKIN-17.  
FT CARBOHYD 68 68 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT SEQUENCE 155 AA; 17504 MW; 2BCAE9CB2F488D01 CRC64;  
SQ  
  
Query Match 3.8%; Score 6; DB 1; Length 155;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 123 VGCTCV 128  
Db 142 VGCTCV 147  
|||||  
  
RESULT 22  
ESTL\_MOUSE STANDARD; PRT; 156 AA.  
AC P11374;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-JUL-1989 (Rel. 11, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE ESTERASE 1 (EC 3.1.1.1) (FRAGMENT).  
GN ESI OR ES-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE-88183471; PubMed-2895647;  
RA Genetta T.L., D'Eustachio P., Kader S.S., Finlay T.H.;  
RT "CDNA cloning of esterase 1, the major esterase activity in mouse  
RL plasma.";  
RL Biochem. Biophys. Res. Commun. 151:1364-1370(1988).  
CC -!- CATALYTIC ACTIVITY: A CARBOXYLIC ESTER + H(2)O = AN ALCOHOL  
CC + A CARBOXYLIC ANION.  
CC -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M19677; AAA37579.1; -  
DR PIR; A27686; A27686.  
DR MGD; MGI:95420; Esl.  
DR InterPro: IPR002018; Carboxylesterase\_B.  
DR PROSITE; PS00122; CARBOXYLESTERASE\_B\_1; PARTIAL.  
DR PROSITE; PS00941; CARBOXYLESTERASE\_B\_2; PARTIAL.  
KW Hydrolase; Serine esterase.  
KW NON\_TER 1  
FT NON\_TER 156 156  
FT SEQUENCE 156 AA; 18360 MW; 39D236E9ECD625C3 CRC64;  
SQ  
  
Query Match 3.8%; Score 6; DB 1; Length 156;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 75 RGCLTG 80  
Db 13 RGCLTG 18  
|||||  
  
RESULT 23  
IL17\_MOUSE STANDARD; PRT; 158 AA.  
ID IL17\_MOUSE  
AC Q62386; Q60971;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED  
DE ANTIGEN 8) (CTLA-8).  
GN IL17 OR CTLA8.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.  
STRAIN-BALB/C; TISSUE-Thymocytes;  
RX MEDLINE-97031826; PubMed-8877732;  
RA Kennedy J., Rossi D.L., Zurawski S.M., Vega F. Jr., Kastelein R.A.,  
Wagner J.L., Hannum C.H., Zlotnik A.;  
RT "Mouse IL-17: a cytokine preferentially expressed by alpha beta TCR +  
RT CD4-CD8-T cells.";  
RL J. Interferon Cytokine Res. 16:611-617(1996).  
[2]  
SEQUENCE OF 12-158 FROM N.A.  
RN STRAIN-129/SV; TISSUE-T-cell;  
RC MEDLINE-96194901; PubMed-8654948;  
RX Yao Z., Timour M., Painter S., Fanslow W., Spriggs M.K.;  
RT "Complete nucleotide sequence of the mouse CTLA8 gene.";  
RL Gene 168:223-225(1996).  
CC -!- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND  
CC HEMATOPOIETIC CYTOKINES (BY SIMILARITY).  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- TISSUE SPECIFICITY: RESTRICTED TO A SUBSET OF ACTIVATED T-CELLS.  
CC -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO  
CC HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U43088; AAB05222.1; -

```
DR EMBL; U35108; AAA93253.1; -.
DR MGD; MGI:107364; I117.
KW Cytokine; Glycoprotein; T-cell; Antigen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 158 INTERLEUKIN-17.
FT CARBOHYD 71 71 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 158 AA; 17490 MW; 3505C143435F4653 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128
Db 145 VGCTCV 150
|||||

RESULT 24
SODC_CAUCR
ID SODC_CAUCR STANDARD; PRT; 174 AA.
AC P20379;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE SUPEROXIDE DISMUTASE [CU-ZN] PRECURSOR (EC 1.15.1.1).
GN SODC OR CC1579.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.K.R., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Shapiro L., Venter J.C., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
RN [3]
RP PARTIAL SEQUENCE.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=82265686; PubMed=7050107;
RA Steinman H.M.;
RT "Copper-zinc superoxide dismutase from Caulobacter crescentus CB15. A
RT novel bacteriocuprein form of the enzyme.";
RL J. Biol. Chem. 257:10283-10293(1982).
CC -1- FUNCTION: MAY FUNCTION AGAINST EXTRACYTOSOLIC TOXIC OXYGEN
CC SPECIES.
CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; M55259; AAA23054.1; -.
CC EMBL; AE005832; AAK23558.1; -.
DR PIR; A35383; A35383.
DR TIGR; P00446; IYAI.
DR TIGR; CC1579; -.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu_1.
DR ProDom: PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc; Signal; Periplasmic; Complete proteome.
FT SIGNAL 1 23
FT CHAIN 24 174 SUPEROXIDE DISMUTASE [CU-ZN].
FT METAL 68 68 COPPER (BY SIMILARITY).
FT METAL 70 70 COPPER (BY SIMILARITY).
FT METAL 86 86 COPPER AND ZINC (BY SIMILARITY).
FT METAL 95 95 ZINC (BY SIMILARITY).
FT METAL 104 104 ZINC (BY SIMILARITY).
FT METAL 107 107 ZINC (BY SIMILARITY).
FT METAL 150 150 COPPER (BY SIMILARITY).
FT DISULFID 175 170 BY SIMILARITY.
SQ SEQUENCE 174 AA; 17100 MW; F3B3C79EF3E3642C CRC64;

Query Match 3.8%; Score 6; DB 1; Length 174;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 AGGRPA 36
Db 133 AGGRPA 138
|||||

RESULT 25
YI3A_LACJO
ID YI3A_LACJO STANDARD; PRT; 177 AA.
AC Q48585;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE INSERTION ELEMENT IS1223 HYPOTHETICAL 20.7 KDA PROTEIN (ORFA).
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11506 / VPI 11088;
RX MEDLINE=94350813; PubMed=8071209;
RA Walker D.C., Klaenhammer T.R.;
RT "Isolation of a novel IS3 group insertion element and construction of
RT an integration vector for Lactobacillus spp.";
RL J. Bacteriol. 176:5330-5340(1994).
CC -1- SIMILARITY: BELONGS TO THE IS150/IS1296 ORFA FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; U09558; AAA56999.1; -.
KW Transposable element; Hypothetical protein.
SQ SEQUENCE 177 AA; 20731 MW; B45019c4BBADCE71 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 177;
Best Local Similarity 100.0%; Pred. No. 1e+02;
```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy ~ 61 ARYPY 66  
| | | | |  
Db 165 ARYPY 170

## RESULT 26

NSG1\_HUMAN STANDARD; PRT; 185 AA.  
AC P42857;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE NEURON SPECIFIC PROTEIN FAMILY MEMBER 1 (BRAIN NEURON CYTOPLASMIC  
DE PROTEIN 1) (D4S234).  
GN NSG1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=97165882; PubMed=9013775;  
RA Carlock L., Vo T., Lorincz M., Walker P.D., Bessert D., Wisniewski D.,  
RA Dunbar J.C.;  
RT "Variable subcellular localization of a neuron-specific protein  
RT during Ntera 2 differentiation into post-mitotic human neurons.";  
RL Brain Res. Mol. Brain Res. 42:202-212(1996).  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE NSG FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M98529; AAA60152.1; -.  
CC EMBL; M98528; AAA60355.1; -.  
CC GDI; MGI:109149; Nsg1.  
KW Golgi stack; Transmembrane; Signal-anchor; Neurone.  
FT DOMAIN 1 82  
FT TRANSMEM 83 103  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT LUMENAL (POTENTIAL).  
FT NV -> KL (IN REF. 1).  
SQ SEQUENCE 185 AA; 20913 MW; 4B7086C18BC11605 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 185;

Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 27

NSG1\_MOUSE STANDARD; PRT; 185 AA.  
AC Q62092; O54717;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE NEURON SPECIFIC PROTEIN FAMILY MEMBER 1 (BRAIN NEURON CYTOPLASMIC  
DE PROTEIN 1) (P21) (M234).  
GN NSG1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;

RN SEQUENCE FROM N.A.  
RP MEDLINE=97165882; PubMed=9013775;  
RX Carlock L., Vo T., Lorincz M., Walker P.D., Bessert D., Wisniewski D.,  
RA Dunbar J.C.;  
RT "Variable subcellular localization of a neuron-specific protein  
RT during Ntera 2 differentiation into post-mitotic human neurons.";  
RL Brain Res. Mol. Brain Res. 42:202-212(1996).  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI (PROBABLE).  
CC -!- TISSUE SPECIFICITY: PITUITARY AND LESS IN ADRENAL GLAND AND  
CC TESTIS.  
CC  
CC -!- SIMILARITY: BELONGS TO THE NSG FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M98530; AAA39966.1; -.  
CC EMBL; AF035683; AAB88210.1; -.  
CC GDI; MGI:109149; Nsg1.  
KW Golgi stack; Transmembrane; Signal-anchor; Neurone.  
FT DOMAIN 1 82  
FT TRANSMEM 83 103  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT LUMENAL (POTENTIAL).  
FT C -> Y (IN REF. 1).  
FT D -> A (IN REF. 2).  
FT NV -> KL (IN REF. 1).  
SQ SEQUENCE 185 AA; 20973 MW; C759938A7171080B CRC64;

Query Match 3.8%; Score 6; DB 1; Length 185;

Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 RSVSPW 52  
| | | | |  
Db 159 RSVSPW 164

## RESULT 28

3MGH\_BORBU STANDARD; PRT; 186 AA.  
ID 3MGH\_BORBU  
AC O51383;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PUTATIVE 3-METHYLADEININE DNA GLYCOSYLASE (EC 3.2.2.-).  
GN BB0422.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35210 / B31;  
RX MEDLINE=98065943; PubMed=9403685;  
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,  
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,

```
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uterback T., Watthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -1- SIMILARITY: BELONGS TO THE MPG FAMILY OF DNA GLYCOSYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001147; AAC66797.1; -.
CC TIGR; BB0422; -.
DR InterPro; IPR003180; Pur_DNA_glyco.
DR Pfam; PF02245; Pur_DNA_glyco; 1.
KW Hypothetical protein; DNA repair; Hydrolase; Complete proteome.
SQ SEQUENCE 186 AA; 21382 MW; 65D80390904825EB CRC64;

Query Match 3.8%; Score 6; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 AKLLLG 152
DB 14 AKLLLG 19
|||||

RESULT 29
UBC4_DROME STANDARD; PRT; 199 AA.
AC P52486;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE UBIQUITIN-CONJUGATING ENZYME E2-22 KDA (EC 6.3.2.19)
DE (UBIQUITIN-PROTEIN LIGASE) (UBIQUITIN CARRIER PROTEIN).
GN UBCD4.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RA Kirby R.J.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE =
CC AMP + PYROPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC STRONGEST, TO YEAST UBC1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X92838; CAA63424.1; -.

```

```
DR HSSP; PI5731; 2UCE.
DR FlyBase; FBgn0015321; Ubcd4.
DR InterPro; IPR000449; UBA.
DR InterPro; IPR000608; UBQ_conjugat.
DR Pfam; PF00627; UBA; 1.
DR Pfam; PF00179; UQ_con; 1.
DR SMART; SM00165; UBA; 1.
DR SMART; SM00212; UBCc; 1.
DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE; PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubiquitin conjugation; Ligase; Multigene family.
FT BINDING 92 92 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 199 AA; 22391 MW; 720CA9595FC0C08F CRC64;

Query Match 3.8%; Score 6; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ARAVLS 10
DB 179 ARAVLS 184
|||||

RESULT 30
HIS5_AQUAE STANDARD; PRT; 207 AA.
AC O66943;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE AMIDOTRANSFERASE HISH (EC 2.4.2.-).
GN HISH OR AQ_732.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: CATALYZES AN AMIDOTRANSFERASE REACTION THAT GENERATES
CC IMIDAZOLE-GLYCEROL PHOSPHATE AND 5-AMINOIMIDAZOL-4-CARBOXAMIDE
CC RIBONUCLEOTIDE, WHICH IS USED FOR PURINE SYNTHESIS.
CC -1- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE HISH FAMILY.
CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000705; AAC06901.1; -.
DR InterPro; IPR000991; GATase_1.
DR Pfam; PF00117; GATase; 1.
DR PROSITE; PS00442; GATASE_TYPE_I; 1.
KW Histidine biosynthesis; Transferase; Glutamine amidotransferase;
KW Complete proteome.
FT ACT_SITE 81 81 BY SIMILARITY.
FT ACT_SITE 183 183 BY SIMILARITY.
FT ACT_SITE 185 185 BY SIMILARITY.
SQ SEQUENCE 207 AA; 23307 MW; 95778F42F098AE38 CRC64;

```

```
Query Match      3.8%; Score 6; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 NLRVSYS 50
      |||||
Db 12 NLRVSYS 17

RESULT 31
UCRL_CHRVI
ID UCRL_CHRVI STANDARD; PRT; 207 AA.
AC 031214;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UBIQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT (EC 1.10.2.2)
DE (RIESKE IRON-SULFUR PROTEIN) (RISP).
GN PETA.
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
OC Allochrotrium
OC NCBI_TaxID=1049;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN=D / ATCC 17899 / DSM 180;
RA Chen Y.L., Dincturk H.B., Qin H., Knaff D.B.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C = Q +
CC 2 FERROCYTOCHROME C.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL).
CC -1- MISCELLANEOUS: THE RIESKE PROTEIN IS A HIGH POTENTIAL 2FE-2S
CC PROTEIN.
CC -1- SIMILARITY: TO RIESKE PROTEINS FROM OTHER SOURCES (MITOCHONDRIA,
CC BACTERIAL, CHLOROPLAST).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF034104; AAB86973.1; -
CC InterPro: IPR001281; Rieske.
CC Pfam: PF00355; Rieske.1.
CC PROSITE; PS00199; RIESKE_1; 1.
CC PROSITE; PS00200; RIESKE_2; 1.
CC Electron transport; Inner membrane; Transmembrane; Iron-sulfur;
CC Oxidoreductase
KW TRANSMEM 24 44 POTENTIAL.
FT METAL 134 134 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 136 136 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 162 162 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 165 165 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT DISULFID 139 164 BY SIMILARITY.
SQ SEQUENCE 207 AA; 22124 MW; C6886EDF25E981CD CRC64;

Query Match      3.8%; Score 6; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SARARA 7
      |||||
```

```
Db 50 SARARA 55

RESULT 32
YOR3_SOUV3
ID YOR3_SOUV3 STANDARD; PRT; 211 AA.
AC Q04550;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 22.3 KDA PROTEIN IN COAT PROTEIN GENE 3'REGION (ORF3).
OS Southampton virus (serotype 3).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OC NCBI_TaxID=37129;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93142023; PubMed=8380940;
RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;
RT "Sequence and genome organization of a human small round-structured
RT (Norwalk-like) virus."
RL Science 259:516-519(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L07418; AAA92985.1; -
CC PIR: C37491; C37491.
CC Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 211 AA; 22257 MW; 3D6D1B51CC3C27B4 CRC64;

Query Match      3.8%; Score 6; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 LRSVSP 51
      |||||
Db 163 LRSVSP 168

RESULT 33
GT29_FASHE
ID GT29_FASHE STANDARD; PRT; 220 AA.
AC P56598;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA 1 (EC 2.5.1.18) (GST1) (FHL) (GST
DE CLASS-ALPHA).
OS Fasciola hepatica (Liver fluke).
OC Rhabdityota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
OC Riaditophora; Eulcithophora; Revertospermata; Mediofusata;
OC Neodermata; Trematoda; Digenea; Echinostomida; Echinostomata;
OC Fascioloides; Fasciolidae; Fasciola.
OC NCBI_TaxID=6192;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92155306; PubMed=1740183;
RA Panaccio M., Wilson L.R., Cramer S.L., Wijffels G.L., Spithill T.W.;
RT "Molecular characterization of cDNA sequences encoding glutathione S-
RT transferases of Fasciola hepatica."
RL Exp. Parasitol. 74:232-237(1992).
RN [2]
RP ERRATUM.
RX MEDLINE=94039664; PubMed=8224094;
RA Panaccio M., Wilson L.R., Cramer S.L., Wijffels G.L., Spithill T.W.;
RL Exp. Parasitol. 77:385-385(1993).
```

[3]  
RN SEQUENCE OF 22-220 FROM N.A.  
RA Cramer1 S.;  
RL Patent number WO9008819, 09-AUG-1990.  
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER  
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.  
CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE  
CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO  
CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF  
CC HAEMATIN IN THE PARASITE GUT.  
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: A00993; CAA00118.1; -;  
CC InterPro: IPR00521; GST.  
CC Pfam: PF00043; GST; 1.  
CC Transferase; Antigen; Multigene family.  
CC INIT\_MET 0 BY SIMILARITY.  
CC CONFLICT 22 22 Y -> V (IN REF. 3).  
CC CONFLICT 110 111 DP -> VS (IN REF. 3).  
CC CONFLICT 189 189 A -> P (IN REF. 3).  
CC SEQUENCE 220 AA; 25598 MW; 27B9F150B75D101F CRC64;  
CC  
Query Match 3.8%; Score 6; DB 1; Length 220;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 55 RISYDP 60  
Db 106 RISYDP 111  
|||||  
RESULT 34  
C1B3\_SHEEP STANDARD; PRT; 232 AA.  
ID C1B3\_SHEEP  
AC P80943;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE T-CELL SURFACE GLYCOPROTEIN CD1B-3 (CD1B-3 ANTIGEN) (SCD1T10)  
DE (FRAGMENT).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Tetal thymocytes;  
RX MEDLINE=96269982; PubMed=8662069;  
RA Ferguson E.E., Dutia B.M., Hein W.R., Hopkins J.;  
RT "The sheep CD1 gene family contains at least four CD1B homologues.";  
RL Immunogenetics 44:86-96(1996).  
CC -1- FUNCTION: NOT KNOWN.  
CC -1- SUBUNIT: ASSOCIATES NON-COVALENTLY WITH BETA-2-MICROGLOBULIN (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X90567; CAA62187.1; -;  
CC DR HSSP; P11609; 1CD1.  
CC DR InterPro: IPR003006; Ig\_MHC.  
CC DR InterPro: IPR003597; Ig\_cl.  
CC Pfam: PF00047; Ig; 1.  
CC DR SMART; SM00407; IGcl; 1.  
CC KW Glycoprotein; Immunoglobulin domain; Transmembrane; Multigene family.  
CC FT NON\_TER 1  
CC FT DOMAIN <1 201 EXTRACELLULAR (POTENTIAL).  
CC FT TRANSMEM 222 232 POTENTIAL.  
CC FT DOMAIN 223 232 CYTOPLASMIC (POTENTIAL).  
CC FT DISULFID 19 83 BY SIMILARITY.  
CC FT DISULFID 123 178 BY SIMILARITY.  
CC FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).  
CC SQ SEQUENCE 232 AA; 26023 MW; C96DB93840B56158 CRC64;  
CC  
Query Match 3.8%; Score 6; DB 1; Length 232;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 127 CVPEPE 132  
Db 48 CVPEPE 53  
|||||  
RESULT 35  
E1A\_ADECT STANDARD; PRT; 232 AA.  
ID E1A\_ADECT  
AC P14264;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE EARLY E1A 25 KDA PROTEIN.  
OS Canine adenovirus type 2 (strain Toronto A 26-61).  
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
OX NCBI\_TaxID=69152;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90021176; PubMed=2800332;  
RA Shibata R., Shinagawa M., Iida Y., Tsukiyama T.;  
RT "Nucleotide sequence of E1 region of canine adenovirus type 2.";  
RL Virology 172:460-467(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Campbell J.B., Zhao Y.;  
RT "Complete DNA sequence and genomic organization of canine  
RT adenovirus type 2.";  
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: J04368; AAA42470.1; -;  
CC DR EMBL; U77082; AAB38711.1; -;  
CC DR PIR; A34165; WMADC2.  
CC DR InterPro: IPR003853; Adeno\_E1A.  
CC DR Pfam; PF02703; Adeno\_E1A; 1.  
CC KW Transcription regulation; Early protein.  
CC SQ SEQUENCE 232 AA; 25346 MW; 2DF9C5983CADFC79 CRC64;  
CC  
Query Match 3.8%; Score 6; DB 1; Length 232;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 CVPEPE 132  
|||||  
Db 213 CVPEPE 218

RESULT 36  
YSR3 CAEEL  
ID YSR3 CAEEL STANDARD; PRT; 233 AA.  
AC Q09951;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL 27.0 KDA PROTEIN F59B10.3 IN CHROMOSOME II.  
GN F59B10.3  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Lloyd C.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z48716; CAA88597.1; -  
DR Wormpep; F59B10.3; CE01592.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 22 42 POTENTIAL.  
FT TRANSMEM 161 181 POTENTIAL.  
SQ SEQUENCE 233 AA; 27039 MW; 50510D335FBD0BC0 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 233;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 SYDPAR 62  
|||||  
Db 221 SYDPAR 226

RESULT 37  
RL4 METH  
ID RL4 METH STANDARD; PRT; 254 AA.  
AC O26111;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE 505 RIBOSOMAL PROTEIN L4/L1E.  
GN RPL4E OR MTH3.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanobacter.  
OX NCBI\_TaxID=145262;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DELTA H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
RA Aldredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,

Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
McDougal S., Shimer G., Goyal A., Pietrovski S., Church G.M.,  
Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
deltaH: functional analysis and comparative genomics.";  
RL J. Bacteriol. 179:7135-7155(1997).  
CC -1- SIMILARITY: BELONGS TO THE L4E FAMILY OF RIBOSOMAL PROTEINS.

CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE000795; AAB84523.1; -  
DR InterPro; IPR002136; Ribosomal\_L4/L1E.  
DR Pfam; PF00573; Ribosomal\_L4; 1.  
DR PROSITE; PS00939; RIBOSOMAL\_L1E; 1.  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 254 AA; 28387 MW; B0EC2B95FE69E4D9 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 254;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RAVLSA 11  
|||||  
Db 33 RAVLSA 38

RESULT 38  
BPHF RHOSO  
ID BPHF RHOSO STANDARD; PRT; 258 AA.  
AC O05151;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE 4-HYDROXY-2-OXOVALERATE ALDOLASE (EC 4.1.2.-).  
GN BPHF OR ETBF.  
OS Rhodococcus sp. (strain RHAL).  
OC Plasmid.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.  
OX NCBI\_TaxID=1831;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97225808; PubMed=9073078;  
RA Masai E., Sugiyama K., Iwashita N., Shimizu S., Hauschild J.E.,  
RA Hatta T., Kimbara K., Yano K., Fukuda M.;  
RT "The bphDEF meta-cleavage pathway genes involved in  
RT biphenyl/polychlorinated biphenyl degradation are located on a linear  
RT plasmid and separated from the initial bphACB genes in Rhodococcus sp.  
RT strain RHAL".  
RL Gene 187:141-149(1997).  
CC -1- FUNCTION: CATALYZES THE PRODUCTION OF PYRUVATE AND ACETALDEHYDE  
CC FROM 4-HYDROXY-2-OXOVALERATE.  
CC -1- PATHWAY: SIXTH STEP IN THE META-CLEAVAGE OF POLYCHLOROBIPHENYLS  
CC (PCB) PATHWAY.  
CC -1- INDUCTION: BY GROWTH ON ETHYLBENZENE OR BIPHENYL.  
CC -1- SIMILARITY: BELONGS TO THE HPHC/HPAI ALDOLASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D78322; BAA18937.1; -

KW Aromatic hydrocarbons catabolism; Lyase; Plasmid.  
SQ SEQUENCE 258 AA; 27159 MW; A539C46C4DA0190F CRC64;

Query Match 3.8%; Score 6; DB 1; Length 258;  
Best Local Similarity 100.0%; Pred. No. 1.4e-02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 RYPRYL 67  
| | | | |  
Db 132 RYPRYL 137

RESULT 39

ID YHEB\_CHLVI STANDARD; PRT; 261 AA.  
AC P56160;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE HYPOTHEICAL 28.2 KDA PROTEIN IN HEMB 3'REGION.  
OS Chlorobium vibrioforme.  
OC Bacteria; Green sulfur bacteria; Chlorobium.  
OX NCBI\_TaxID=1098;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=F. THIOSULFATOPHILUM / NCIB 8327;  
RX MEDLINE=96215215; PubMed=8626508;  
RA Rhie G.-E., Avissar Y.J., Beale S.I.;  
RT "Structure and expression of the Chlorobium vibrioforme hemb gene and  
characterization of its encoded enzyme, porphobilinogen synthase.";  
RL J. Biol. Chem. 271:8176-8182(1996).  
CC -!- SIMILARITY: BELONGS TO THE INOSITOL MONOPHOSPHATASE FAMILY  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U38348; -; NOT\_ANNOTATED\_CDS.  
DR HSSP; P29218; 2HHM.  
DR InterPro; IPR000760; Inositol\_P.  
DR Pfam; PF00459; Inositol\_P\_2.  
DR PRINTS; PR00378; INOSPHPTASE.  
DR PROSITE; PS00629; IMP\_1; 1.  
DR PROSITE; PS00630; IMP\_2; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 261 AA; 28243 MW; D25E7D223024095F CRC64;

Query Match 3.8%; Score 6; DB 1; Length 261;  
Best Local Similarity 100.0%; Pred. No. 1.4e-02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GLFCEE 85  
| | | | |  
Db 64 GLFCEE 69

RESULT 40

ID OTCC\_NEICI STANDARD; PRT; 262 AA.  
AC Q01322;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE ORNITHINE CARBAMOYLTRANSFERASE, CATABOLIC (EC 2.1.3.3) (OTCASE)  
DE (FRAGMENT).  
GN ARGF.  
OS Neisseria cinerea.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=483;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LNP 1646;  
RX MEDLINE=93023840; PubMed=1406254;  
RA Zhou J., Spratt B.G.;  
RT "Sequence diversity within the argF, fbp and recA genes of natural  
isolates of Neisseria meningitidis: interspecies recombination within  
the argF gene.";  
RL Mol. Microbiol. 6:2135-2146(1992).  
CC -!- CATALYTIC ACTIVITY: CARBAMOYL PHOSPHATE + ORNITHINE = CITRULLINE  
CC + ORTHOPHOSPHATE.  
CC -!- PATHWAY: SECOND STEP IN ARGININE DEGRADATION VIA ARGININE  
CC DEIMINASE.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X64869; CAA46081.1; -.  
DR PIR; S24718; S24718.  
DR HSSP; P08308; IORT.  
DR InterPro; IPR002029; Carbmyltransf\_asor.  
DR Pfam; PF00185; OTCase; 1.  
DR Pfam; PF02729; OTCase.N; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; PARTIAL.  
KW Transferase; Arginine metabolism.  
FT NON\_TER 1  
FT NON\_TER 262  
SQ SEQUENCE 262 AA; 29037 MW; 589D1721FD5E8B4A CRC64;

Query Match 3.8%; Score 6; DB 1; Length 262;  
Best Local Similarity 100.0%; Pred. No. 1.4e-02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ARARAV 8  
| | | | |  
Db 146 ARARAV 151

RESULT 41

ID OTCC\_NEIMU STANDARD; PRT; 262 AA.  
AC Q01326;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE ORNITHINE CARBAMOYLTRANSFERASE, CATABOLIC (EC 2.1.3.3) (OTCASE)  
DE (FRAGMENT).  
GN ARGF.  
OS Neisseria mucosa.  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LNP 405;  
RX MEDLINE=93023840; PubMed=1406254;  
RA Zhou J., Spratt B.G.;  
RT "Sequence diversity within the argF, fbp and recA genes of natural  
isolates of Neisseria meningitidis: interspecies recombination within  
the argF gene.";  
RL Mol. Microbiol. 6:2135-2146(1992).  
CC -!- CATALYTIC ACTIVITY: CARBAMOYL PHOSPHATE + ORNITHINE = CITRULLINE  
CC + ORTHOPHOSPHATE.  
CC -!- PATHWAY: SECOND STEP IN ARGININE DEGRADATION VIA ARGININE

```

CC DEIMINASE.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X64873; CAA46085.1; -
CC DR PIR; S24727; S24727.
CC DR HSSP; P08308; IORT.
CC DR InterPro; IPR002029; Carbmyltransf_asor.
CC DR Pfam; PF00185; OTCace.1.
CC DR Pfam; PF02729; OTCace.N.1.
CC DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; PARTIAL.
CC KW Transferase; Arginine metabolism.
CC FT NON_TER 1 1
CC FT 262 262
CC SEQUENCE 262 AA; 29063 MW; BDE7B8F8D1CC48B4 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ARARAV 8
DB 146 ARARAV 151
|||||

RESULT 42
FTSQ_STRCO STANDARD; PRT; 264 AA.
AC P45518;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION PROTEIN FTSQ HOMOLOG.
GN FTSQ OR SC4A10.16C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=96359387; PubMed=8752351;
RA McCormick J.R., Losick R.;
RT "Cell division gene ftsQ is required for efficient sporulation but
RT not growth and viability in Streptomyces coelicolor A3(2).";
RL J. Bacteriol. 178:5295-5301(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C., Harris D., James K.D., Parkhill J., Barrell B.G.,
RA Rajandream M.A.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 236-264 FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=95317146; PubMed=7830569;
RA McCormick J.R., Su E.P., Driks A., Losick R.;
RT "Growth and viability of Streptomyces coelicolor mutant for the cell
RT division gene ftsZ.";
RL Mol. Microbiol. 14:243-254(1994).
CC -!- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN SEPTUM FORMATION (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE FTSQ FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U10879; AAD10532.1; -
CC DR EMBL; ALI09663; CAB51992.1; -
CC KW Cell division; Septation; Transmembrane.
CC FT DOMAIN 1 32 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 33 53 POTENTIAL.
CC FT DOMAIN 54 264 EXTRACELLULAR (POTENTIAL).
CC SEQUENCE 264 AA; 28655 MW; 72FE8048C62BFCEA CRC64;

Query Match 3.8%; Score 6; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 RFRFP 43
DB 23 RFRFP 28
|||||

RESULT 43
YEDP_ECOLI STANDARD; PRT; 271 AA.
ID YEDP_ECOLI
AC P76329;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 30.4 KDA PROTEIN IN DSRB-VSR INTERGENIC REGION.
GN YEDP OR B1955.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000287; AAC75021.1; -
CC DR EcoGene; EGI4039; yedP.
CC DR InterPro; IPR001454; Hydrolase.
CC DR Pfam; PF00702; Hydrolase.1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 271 AA; 30439 MW; 1192A5DC83F71816 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 GPNDAP 157
|||||

```

```

Db      215 GPNDAP 220

RESULT 44
LIP_THELA
ID LIP_THELA STANDARD; PRT; 291 AA.
AC O59952;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LIPASE PRECURSOR (EC 3.1.1.3) (TRIACYLGLYCEROL LIPASE).
GN LIP.
OS Thermomyces lanuginosus (Humicola lanuginosa).
OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Thermomyces.
OX NCBI_TaxID=5541;
RN [1]
RP SEQUENCE FROM N.A.
RA Boel E., Muller S., Sandal T., Kamp-Hansen P., Dalboge H.;
RT "Wild type Humicola lanuginosa cDNA encoding a lipolytic enzyme.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP MUTAGENESIS OF TRP-111.
RX MEDLINE=95115502; PubMed=7815893;
RA Holmquist M., Martinelle M., Clausen I.G., Patkar S., Svendsen A.,
RA Hult K.;
RT "Trp89 in the lid of Humicola lanuginosa lipase is important for
RT efficient hydrolysis of tributyrin.";
RL Lipids 29:599-603(1994).
RN [3]
RP MUTAGENESIS
RX MEDLINE=95242058; PubMed=7536956;
RA Naver H., Lovborg U.;
RT "The importance of non-charged amino acids in antibody binding to
RT Humicola lanuginosa lipase.";
RL Scand. J. Immunol. 41:443-448(1995).
RN [4]
RP MUTAGENESIS OF GLU-109 AND TRP-111.
RX MEDLINE=97015915; PubMed=8862552;
RA Martinelle M., Holmquist M., Clausen I.G., Patkar S., Svendsen A.,
RA Hult K.;
RT "The role of Glu87 and Trp89 in the lid of Humicola lanuginosa
RT lipase.";
RL Protein Eng. 9:519-524(1996).
RN [5]
RP MUTAGENESIS OF SER-168.
RX MEDLINE=98400942; PubMed=9730809;
RA Peters G.H., Svendsen A., Langberg H., Vind J., Patkar S.A.,
RA Toxvaerd S., Kinnunen P.K.;
RT "Active serine involved in the stabilization of the active site loop
RT in the Humicola lanuginosa lipase.";
RL Biochemistry 37:12375-12383(1998).
CC -1- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL +
CC A FATTY ACID ANION.
CC -1- BIOTECHNOLOGY: USED AS A DETERGENT LIPASE. SOLD UNDER THE NAME
CC LIPOLEASE BY NOVOTYMS. ENGINEERED VARIANTS ARE SOLD UNDER THE
CC NAMES LIPOLEASE ULTRA AND LIPOPRIME.
CC -1- MISCELLANEOUS: EFFECTIVE UNDER ALKALINE CONDITIONS (UP TO PH 12
CC APPROXIMATELY) AND AT A BROAD TEMPERATURE RANGE.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF054513; AAC08588.1;
DR HSSP; P21811; 1T1B.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR002921; Lipase_3.

```

---

```

DR Pfam; PF01764; Lipase_3; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Zymogen; Signal.
FT SIGNAL 1 17
FT PROPEP 18 22
FT CHAIN 23 291
FT ACT_SITE 168 168
FT ACT_SITE 223 223
FT ACT_SITE 280 280
SQ SEQUENCE 291 AA; 31806 MW; 170ACEDF791DB07B CRC64;

Query Match 3.8%; Score 6; DB 1; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159
DB 48 NDAPAG 53
|||||

RESULT 45
YX95_MYCTU STANDARD; PRT; 294 AA.
ID YX95_MYCTU
AC Q50730; O50420;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE HYPOTHETICAL 29.9 KDA PROTEIN RV3395C.
GN RV3395C OR MT3502 OR MTC178.33.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AL009198; CAA15780.1;
DR EMBL; AE007156; AAK47839.1; ALT_INIT.
DR TIGR; MT3502;
DR TuberculList; Rv3395C;

```

KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 153 173 POTENTIAL.  
 SQ SEQUENCE 294 AA; 29904 MW; 7E920C48A5ED096B CRC64;

Query Match 3.8%; Score 6; DB 1; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ARARAV 8  
 |||||  
 Db 228 ARARAV 233

RESULT 46  
 MPRL SCHPO STANDARD; PRT; 295 AA.  
 AC 094321; (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE MULTISTEP PHOSPHORELAY REGULATOR 1.  
 GN MPRL OR SPEC725.02.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., MUTAGENESIS OF HIS-221, AND FUNCTION.  
 RX MEDLINE=20214813; PubMed=10749922;  
 RA Nguyen A.N., Lee A., Place W., Shiozaki K.;  
 RT "Multistep phosphorelay proteins transmit oxidative stress  
 signals to the fission yeast stress-activated protein kinase."  
 RL Mol. Biol. Cell 11:1169-1181(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Lyne M., Rajandream M.A., Barrell B.G., Rieger M.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: BINDS TO THE MSC4 RESPONSE REGULATOR WHICH IS PART OF A  
 CC MULTISTEP PHOSPHORELAY SYSTEM THAT TRANSMITS OXIDATIVE STRESS  
 CC SIGNALS TO THE SPC1 MAPK CASCADE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL034352; CAA22174.1; -  
 CC InterPro; IPR002570; Hpt.  
 CC SMART; SM00073; HPT; 1.  
 KW Phosphorylation.  
 FT MOD\_RES 221 221 PHOSPHORYLATION (POTENTIAL).  
 FT MUTAGEN 221 221 H->Q: OXIDATIVE STRESS SIGNALING TO  
 FT SPC1 DID NOT OCCUR.  
 SQ SEQUENCE 295 AA; 32550 MW; 99207DA474E4F9E9 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 QARNAS 28  
 |||||  
 Db 26 QARNAS 31

RESULT 47  
 YBHK\_ECOLI

ID YBHK\_ECOLI STANDARD; PRT; 302 AA.  
 AC P75767;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DE 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE HYPOTHETICAL PROTEIN YBHK.  
 GN YBHK OR B0780.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12."  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horiuchi T.;  
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome  
 corresponding to the 12.7-28.0 min region on the linkage map."  
 RL DNA Res. 3:137-155(1996).  
 CC -1- SIMILARITY: BELONGS TO THE UPF0052 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE000180; AAC73867.1; -  
 CC EMBL; D90716; BAA35438.1; -  
 CC EcoGene; EGI3667; ybhK.  
 CC InterPro; IPR002882; UPF0052.  
 CC Pfam; PF01933; UPF0052; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 302 AA; 32788 MW; 4E88D67B89576874 CRC64;

Query Match 3.8%; Score 6; DB 1; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 LRRTPA 106  
 |||||  
 Db 210 LRRTPA 215

RESULT 48  
 SECF\_RICPR STANDARD; PRT; 311 AA.  
 ID SECF\_RICPR  
 AC Q9ZE34;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROTEIN-EXPORT MEMBRANE PROTEIN SECF.  
 GN SECF OR RP114.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 OX NCBI\_TaxID=782;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANTS Y-74; T-111; V-146 AND T-218.  
 RC STRAIN-MADRID B.  
 RX MEDLINE=99039499; PubMed=9823893;  
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
 RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RT mitochondria";  
 RL Nature 396:133-140(1998).  
 CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT.  
 CC -1- SUBUNIT: PART OF THE PROKARYOTIC PROTEIN TRANSLLOCATION APPARATUS  
 CC WHICH COMPRISE SECA, SECB, SECD, SECE, SECF, SECG AND SECH  
 CC (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE SECD/SECF FAMILY. SECF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: AJ235270; CAA14583.1; -  
 DR InterPro: IPR003335; SecD\_SecF; 1.  
 DR Pfam: PF02355; SecD\_SecF; 1.  
 DR Complete proteome.  
 KW Protein transport; Translocation; Transmembrane; Membrane;  
 FT TRANSMEM 36 56 POTENTIAL.  
 FT TRANSMEM 140 160 POTENTIAL.  
 FT TRANSMEM 164 184 POTENTIAL.  
 FT TRANSMEM 194 214 POTENTIAL.  
 FT TRANSMEM 246 266 POTENTIAL.  
 FT TRANSMEM 272 292 POTENTIAL.  
 SQ SEQUENCE 311 AA; 35175 MW; 31863A150F32E630 CRC64;  
 -----  
 Query Match 3.8%; Score 6; DB 1; Length 311;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 RAVLSA 11  
 Db 70 RAVLSA 75  
 -----  
 RESULT 49  
 O2J2\_HUMAN STANDARD; PRT; 312 AA.  
 AC 076002; O9GZK2; O9GZL3;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE OLFACTORY RECEPTOR 2J2 (OLFACTORY RECEPTOR 6-8) (OR6-8) (HS6MI-6).  
 GN OR2J2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Evans K.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANTS Y-74; T-111; V-146 AND T-218.  
 RA Ziegler A., Ehlers A., Forbes S.A., Trowsdale J., Uchanska-Ziegler B.,  
 RA Volz A., Younger R., Beck S.;  
 RT "Polymorphic olfactory receptor genes and HLA loci constitute extended  
 RT haplotypes";

(In) Kasahara M. (eds.);  
 RL Major histocompatibility complex-evolution, structure, and function,  
 RL pp.110-130. Springer-Verlag, Tokyo (2000).  
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: AL023727; CAA18784.1; -  
 DR EMBL: AJ302571; CAC20491.1; -  
 DR EMBL: AJ302572; CAC20492.1; -  
 DR EMBL: AJ302573; CAC20493.1; -  
 DR EMBL: AJ302574; CAC20494.1; -  
 DR EMBL: AJ302575; CAC20495.1; -  
 DR EMBL: AJ302576; CAC20496.1; -  
 DR EMBL: AJ302577; CAC20497.1; -  
 DR EMBL: AJ302578; CAC20498.1; -  
 DR EMBL: AJ302579; CAC20499.1; -  
 DR EMBL: AJ302580; CAC20500.1; -  
 DR EMBL: AJ302581; CAC20501.1; -  
 DR EMBL: AJ302582; CAC20502.1; -  
 DR EMBL: AJ302583; CAC20503.1; -  
 DR InterPro: IPR000276; GPCR\_Rhodopsin.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PROSITE: PS00237; G-PROTEIN\_RECEPTOR\_F1\_1; 1.  
 DR PROSITE: PS0262; G-PROTEIN\_RECEPTOR\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Olfaction; Polymorphism;  
 FT DOMAIN 1 26 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 27 50  
 FT DOMAIN 51 58 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 59 80  
 FT DOMAIN 81 101 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 102 121  
 FT DOMAIN 122 140 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 141 159  
 FT DOMAIN 160 196 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 197 220  
 FT DOMAIN 221 237 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 238 260  
 FT DOMAIN 261 273 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 274 293  
 FT DOMAIN 294 312 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 98 190  
 FT CARBOHYD 6 6  
 FT VARIANT 74 74  
 FT VARIANT 111 111  
 FT VARIANT 146 146  
 FT VARIANT 218 218  
 FT VARIANT 218 218  
 SQ SEQUENCE 312 AA; 35204 MW; FB7D20BB2379C43E CRC64;  
 -----  
 Query Match 3.8%; Score 6; DB 1; Length 312;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 ARAVLS 10  
 Db 223 ARAVLS 228  
 -----  
 RESULT 50

```

CASB_HUMAN
ID CASB_HUMAN STANDARD; RET; 317 AA.
AC 09Y2D0;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CARBONIC ANHYDRASE VB, MITOCHONDRIAL PRECURSOR (EC 4.2.1.1) (CARBONATE
DE DEHYDRATASE VB) (CA-VB).
GN CASB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Pancreas;
RX MEDLINE=99340061; PubMed=10409679;
RA Fujikawa-Adachi K., Nishimori I., Taguchi T., Onishi S.;
RT "Human mitochondrial carbonic anhydrase VB: cDNA cloning, mRNA
RL expression, subcellular localization, and mapping to chromosome X.";
RL J. Biol. Chem. 274:21228-21233(1999).
CC -!- FUNCTION: REVERSIBLE HYDRATATION OF CARBON DIOXIDE.
CC -!- CATALYTIC ACTIVITY: H(2)CO(3) -> CO(2) + H(2)O.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -!- TISSUE SPECIFICITY: STRONGEST EXPRESSION IN HEART, PANCREAS,
CC KIDNEY, PLACENTA, LUNG, AND SKELETAL MUSCLE. NOT EXPRESSED IN
CC LIVER.
CC -!- SIMILARITY: BELONGS TO THE EUKARYOTIC-TYPE CARBONIC ANHYDRASE
CC FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB021660; BAA76671.1; -.
CC MIM; 300230; -.
DR InterPro; IPR001148; Carb_anhydrase.
DR Pfam; PF00194; carb_anhydrase; 1.
DR ProDom; PD000865; Carb_anhydrase; 1.
DR PROSITE; PS00162; EUK_CO2_ANHYDRASE; 1.
KW Lyase; Zinc; Mitochondrion; Transit peptide.
FT TRANSIT 1 33 MITOCHONDRION.
FT CHAIN 34 317 CARBONIC ANHYDRASE VB.
FT METAL 130 130 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 132 132 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 155 155 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 317 AA; 36433 MW; 7C11920EFF2588A CRC64;

```

Query Match 3.8%; Score 6; DB 1; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 ISYDPA 61  
 |||||  
 Db 85 ISYDPA 90

Search completed: April 14, 2002, 12:45:54  
 Job time: 201 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2002, 12:42:18 ; Search time 25.05 Seconds  
(without alignments)  
934.274 Million cell updates/sec

Title: US-09-731-816-4

Perfect score: 160

Sequence: 1 NSGARAVLSAFHHTLQLGP.....SIDKQAKLLIGPNDAPAG 160

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 6

Total number of hits satisfying chosen parameters: 846

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL\_17.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	5.6	477	4 Q9H7L3	Q9h7l3 homo sapien
2	9	5.6	540	4 Q9H7Q8	Q9h7q8 homo sapien
3	9	5.6	672	4 Q9H7I6	Q9h7i6 homo sapien
4	9	5.6	772	4 Q9H7Q7	Q9h7q7 homo sapien
5	9	5.6	913	11 Q99NC8	Q99nc8 rattus norv
6	9	5.6	914	4 Q9NQR5	Q9nqr5 homo sapien
7	9	5.6	914	4 Q9BYI0	Q9byi0 homo sapien
8	9	5.6	914	4 Q9BXP2	Q9bxp2 homo sapien
9	9	5.6	914	11 Q99WR3	Q99wr3 mus musculu
10	8	5.0	261	2 Q9HX52	Q9hx52 pseudomonas
11	8	5.0	804	4 Q9URZ6	Q9urz6 homo sapien
12	8	5.0	805	4 Q9NRA7	Q9nra7 homo sapien
13	8	5.0	805	4 Q9BYF1	Q9byf1 homo sapien
14	7	4.4	47	7 Q9GIW8	Q9giw8 monachus sc
15	7	4.4	56	2 Q9PG71	Q9pg71 xylella fas
16	7	4.4	92	2 Q9KW34	Q9kw34 pseudomonas
17	7	4.4	131	5 Q9VX98	Q9vx98 drosophila
18	7	4.4	150	1 Q9YCT7	Q9yct7 aeropyrum p
19	7	4.4	154	1 Q29550	Q29550 archaeoglob

20	7	4.4	154	2 Q9RFQ8	Q9rfq8 rhodothermu
21	7	4.4	170	2 Q26071	Q26071 helicobacte
22	7	4.4	170	2 Q9ZJ62	Q9zj62 helicobacte
23	7	4.4	197	4 Q9P0M4	Q9p0m4 homo sapien
24	7	4.4	197	4 Q9HC75	Q9hc75 homo sapien
25	7	4.4	209	2 Q9S558	Q9s558 pseudomonas
26	7	4.4	209	2 Q9S566	Q9s566 pseudomonas
27	7	4.4	243	2 Q96378	Q96378 mycobacteri
28	7	4.4	253	2 Q9K3H8	Q9k3h8 streptomyce
29	7	4.4	288	2 Q50480	Q50480 streptomyce
30	7	4.4	293	2 Q96356	Q96356 mycobacteri
31	7	4.4	307	2 Q26103	Q26103 helicobacte
32	7	4.4	307	2 Q9ZJ28	Q9zj28 helicobacte
33	7	4.4	331	2 Q9ZJ28	Q9zj28 helicobacte
34	7	4.4	332	2 Q9ZJ34	Q9zj34 rhodobacter
35	7	4.4	367	12 Q9Z538	Q9z538 streptomyce
36	7	4.4	399	5 Q9YUR7	Q9yur7 turkey aden
37	7	4.4	435	5 Q16765	Q16765 caenorhabdi
38	7	4.4	467	2 Q9NFY2	Q9nfy2 anopheles g
39	7	4.4	467	2 Q51006	Q51006 neisseria g
40	7	4.4	467	2 Q9JY68	Q9jy68 neisseria m
41	7	4.4	470	3 Q9JT52	Q9jt52 neisseria m
42	7	4.4	486	2 Q9P8G5	Q9p8g5 neurospora
43	7	4.4	551	4 Q9KHC4	Q9khc4 myxococcos
44	7	4.4	551	4 Q13512	Q13512 homo sapien
45	7	4.4	551	4 Q15740	Q15740 homo sapien
46	7	4.4	699	4 Q9UDU4	Q9udu4 homo sapien
47	7	4.4	812	10 Q9ASG3	Q9asg3 oryza sativ
48	7	4.4	830	4 Q43701	Q43701 homo sapien
49	7	4.4	830	4 Q14162	Q14162 homo sapien
50	7	4.4	946	12 Q69139	Q69139 human herpe
51	7	4.4	966	5 Q44066	Q44066 paraurostyl
52	7	4.4	1088	5 Q00905	Q00905 oxytricha f
53	7	4.4	1180	12 Q84919	Q84919 pepper ring
54	7	4.4	1190	12 Q9P252	Q9p252 turkey herp
55	7	4.4	1490	5 Q22216	Q22216 gallid herp
56	7	4.4	1698	12 Q84918	Q84918 caenorhabdi
57	7	4.4	1783	4 Q95321	Q95321 pepper ring
58	7	4.4	1804	4 Q9UNU4	Q9unu4 homo sapien
59	7	4.4	1822	11 Q35412	Q35412 rattus norv
60	7	4.4	2129	12 Q9J6K9	Q9j6k9 rubella vir
61	7	4.4	3086	12 Q11436	Q11436 ryegrass mo
62	7	4.4	3086	12 Q89525	Q89525 ryegrass mo
63	6	3.8	26	4 Q9UGK1	Q9ugk1 homo sapien
64	6	3.8	37	2 Q30792	Q30792 erwinia amy
65	6	3.8	44	2 Q9X715	Q9x715 prochloroco
66	6	3.8	57	7 Q30002	Q30002 homo sapien
67	6	3.8	58	9 Q9T0T5	Q9t0t5 bacterioph
68	6	3.8	66	2 Q74788	Q74788 synechocyst
69	6	3.8	68	7 Q30121	Q30121 homo sapien
70	6	3.8	73	7 Q95381	Q95381 homo sapien
71	6	3.8	77	5 Q9GUN2	Q9gun2 caenorhabdi
72	6	3.8	78	4 Q9NUE6	Q9nue6 homo sapien
73	6	3.8	79	2 Q55261	Q55261 streptococc
74	6	3.8	80	7 Q9TPA8	Q9tpa8 homo sapien
75	6	3.8	81	6 Q9TSK7	Q9tsk7 macaca fasc
76	6	3.8	81	7 Q30731	Q30731 macaca neme
77	6	3.8	82	2 Q9L084	Q9l084 streptomyce
78	6	3.8	82	2 Q30949	Q30949 pan troglod
79	6	3.8	82	12 Q68844	Q68844 hepatitis c
80	6	3.8	85	5 Q17337	Q17337 caenorhabdi
81	6	3.8	87	12 Q9WEY6	Q9wey6 human callic
82	6	3.8	89	7 Q9BD02	Q9bd02 homo sapien
83	6	3.8	90	2 Q9ZHI3	Q9zhi3 yersinia pe
84	6	3.8	90	7 Q30100	Q30100 homo sapien
85	6	3.8	97	12 Q9YZB0	Q9yzb0 human callic
86	6	3.8	97	12 Q9WMA5	Q9wma5 human callic
87	6	3.8	97	12 Q9WB21	Q9wb21 human callic
88	6	3.8	99	12 Q9PE5	Q9pe5 norwalk-lik
89	6	3.8	105	11 Q9D286	Q9d286 mus musculu
90	6	3.8	105	11 Q99P86	Q99p86 mus musculu
91	6	3.8	106	2 Q9XAG6	Q9xag6 streptomyce
92	6	3.8	106	2 Q9XAG6	Q9xag6 streptomyce

93	6	3.8	106	8	P92546	P92546 arabidopsis	166	6	3.8	199	1	Q9C4N8	Q9c4n8 archaeoglob
94	6	3.8	106	10	P93299	P93299 arabidopsis	167	6	3.8	199	2	Q9RMF4	Q9rmf4 deinococcus
95	6	3.8	107	2	Q9F169	Q9f169 salmonella	168	6	3.8	199	2	Q9JMU8	Q9jmu8 buchnera ap
96	6	3.8	108	12	Q98225	Q98225 molluscum c	169	6	3.8	199	4	Q9H7N0	Q9h7n0 homo sapien
97	6	3.8	111	10	Q9ARM9	Q9arm9 oryza sativ	170	6	3.8	199	5	P91633	P91633 drosophila
98	6	3.8	111	11	Q9EP95	Q9ep95 mus musculu	171	6	3.8	200	9	Q38189	Q38189 bacterioph
99	6	3.8	111	11	Q99P85	Q99p85 rattus norv	172	6	3.8	201	5	Q26426	Q26426 sepi
100	6	3.8	113	2	Q54773	Q54773 synechococ	173	6	3.8	202	11	Q62508	Q62508 mus musculu
101	6	3.8	114	2	Q9S206	Q9s206 streptomyce	174	6	3.8	202	11	Q62084	Q62084 mus musculu
102	6	3.8	114	2	Q9PI53	Q9pi53 campylobact	175	6	3.8	204	2	Q9KLA0	Q9kla0 vibrio chol
103	6	3.8	116	12	Q99224	Q99224 human calic	176	6	3.8	204	2	Q9A580	Q9a580 caulobacter
104	6	3.8	117	11	Q908Y4	Q9d8y4 mus musculu	177	6	3.8	205	5	Q9N7M8	Q9n7m8 leishmania
105	6	3.8	118	1	Q9Y951	Q9y951 aeropyrum p	178	6	3.8	205	10	Q80363	Q80363 nicotiana t
106	6	3.8	118	5	Q9N716	Q9n716 leishmania	179	6	3.8	207	12	Q91028	Q91028 human immun
107	6	3.8	119	5	Q9XVR3	Q9xvr3 boophilus m	180	6	3.8	208	3	Q9Y7T9	Q9y7t9 schizosacch
108	6	3.8	120	13	Q9PT93	Q9pt93 brachydanio	181	6	3.8	208	9	Q38188	Q38188 bacterioph
109	6	3.8	128	5	Q9N557	Q9n557 caenorhabdi	182	6	3.8	215	2	Q06169	Q06169 mycobacteri
110	6	3.8	131	5	Q9XWL1	Q9xwl1 caenorhabdi	183	6	3.8	215	2	Q9X7B4	Q9x7b4 mycobacteri
111	6	3.8	133	12	Q9J5E3	Q9j5e3 fowlpox vir	184	6	3.8	217	2	P72402	P72402 streptomyce
112	6	3.8	136	2	Q9RWD3	Q9rwd3 deinococcus	185	6	3.8	218	4	Q9HB02	Q9hb02 homo sapien
113	6	3.8	136	10	P93092	P93092 casuarina g	186	6	3.8	218	13	Q90764	Q90764 gallus gall
114	6	3.8	139	2	Q9RF35	Q9rf35 tolypotbrix	187	6	3.8	219	10	Q9SYN5	Q9syn5 arabidopsis
115	6	3.8	139	2	Q9RAW4	Q9raw4 geitlerinem	188	6	3.8	219	13	Q90575	Q90575 gallus gall
116	6	3.8	139	2	Q9R6W4	Q9r6w4 nostoc sp.	189	6	3.8	220	2	Q9CMS8	Q9cms8 pasteurella
117	6	3.8	142	12	Q9YZB2	Q9yzb2 human calic	190	6	3.8	222	5	Q23183	Q23183 caenorhabdi
118	6	3.8	142	12	Q9YIU9	Q9yiu9 human calic	191	6	3.8	224	2	Q47391	Q47391 escherichia
119	6	3.8	143	2	Q9AN81	Q9an81 bradyrhizob	192	6	3.8	226	9	Q38190	Q38190 bacterioph
120	6	3.8	144	2	Q9R6W3	Q9r6w3 pseudanabae	193	6	3.8	227	5	Q9BJC9	Q9bjc9 brugia mala
121	6	3.8	144	12	Q9YZD3	Q9yzd3 human calic	194	6	3.8	229	7	Q30081	Q30081 homo sapien
122	6	3.8	144	12	Q9YJ87	Q9yj87 human calic	195	6	3.8	232	2	Q08385	Q08385 neisseria f
123	6	3.8	145	2	Q9RF33	Q9rf33 pseudanabae	196	6	3.8	232	2	Q08640	Q08640 neisseria p
124	6	3.8	147	10	Q9AWH8	Q9awh8 oryza sativ	197	6	3.8	232	2	Q08642	Q08642 neisseria s
125	6	3.8	149	2	P71648	P71648 mycobacteri	198	6	3.8	232	2	Q08141	Q08141 neisseria c
126	6	3.8	149	10	Q9AX55	Q9ax55 oryza sativ	199	6	3.8	234	5	Q9VM06	Q9vm06 drosophila
127	6	3.8	151	10	Q9WB30	Q9wb30 pharbitis n	200	6	3.8	235	2	Q9X5R2	Q9x5r2 streptomyce
128	6	3.8	151	12	Q40633	Q40633 salmirline	201	6	3.8	235	10	Q9AS21	Q9as21 oryza sativ
129	6	3.8	154	6	Q9GME7	Q9gme7 pan troglod	202	6	3.8	237	2	Q08353	Q08353 mycobacteri
130	6	3.8	154	6	Q9GME6	Q9gme6 pan troglod	203	6	3.8	237	7	Q30098	Q30098 homo sapien
131	6	3.8	156	2	Q9JUG8	Q9jug8 neisseria m	204	6	3.8	238	2	Q9S2T0	Q9s2t0 streptomyce
132	6	3.8	160	1	Q29917	Q29917 archaeoglob	205	6	3.8	238	2	Q9F5V1	Q9f5v1 azotobacter
133	6	3.8	160	5	Q9GWK0	Q9gwk0 leishmania	206	6	3.8	242	6	Q28075	Q28075 bos taurus
134	6	3.8	160	10	Q41861	Q41861 zea mays (m	207	6	3.8	243	13	Q919T9	Q919t9 lagopus leu
135	6	3.8	165	2	Q9PHN0	Q9phn0 campylobact	208	6	3.8	244	2	Q51345	Q51345 pseudomonas
136	6	3.8	165	2	Q9JRR8	Q9jrr8 neisseria m	209	6	3.8	246	2	Q08376	Q08376 treponema p
137	6	3.8	166	5	Q27051	Q27051 trypanosoma	210	6	3.8	247	5	Q9VMF8	Q9vmf8 drosophila
138	6	3.8	167	11	Q9CUN1	Q9cun1 mus musculu	211	6	3.8	247	5	Q9VEC2	Q9vec2 drosophila
139	6	3.8	172	10	Q9FU90	Q9fu90 oryza sativ	212	6	3.8	249	3	Q9P7T7	Q9p7t7 schizosacch
140	6	3.8	175	2	Q9HX29	Q9hxx29 pseudomonas	213	6	3.8	250	4	Q9NSV7	Q9nsv7 homo sapien
141	6	3.8	176	2	Q9RPS9	Q9rps9 enterococcu	214	6	3.8	255	7	Q19506	Q19506 homo sapien
142	6	3.8	180	1	Q9HMF4	Q9hmf4 halobacteri	215	6	3.8	257	2	Q9ABD3	Q9abd3 caulobacter
143	6	3.8	180	7	Q9GJ09	Q9gj09 homo sapien	216	6	3.8	257	12	Q9YMS7	Q9yms7 lymantria d
144	6	3.8	180	7	Q9GJ07	Q9gj07 homo sapien	217	6	3.8	257	12	Q9YMS7	Q9yms7 lymantria d
145	6	3.8	180	7	Q9GJ05	Q9gj05 homo sapien	218	6	3.8	260	1	Q59348	Q59348 pyrococcus
146	6	3.8	180	7	Q9GJ02	Q9gj02 homo sapien	219	6	3.8	260	1	Q9V1B1	Q9v1b1 pyrococcus
147	6	3.8	180	7	Q9G122	Q9g122 homo sapien	220	6	3.8	261	10	Q9CAE4	Q9cae4 arabidopsis
148	6	3.8	181	7	P78551	P78551 homo sapien	221	6	3.8	263	5	Q27058	Q27058 trypanosoma
149	6	3.8	181	7	Q99693	Q99693 homo sapien	222	6	3.8	263	5	Q27059	Q27059 trypanosoma
150	6	3.8	181	7	Q99694	Q99694 homo sapien	223	6	3.8	264	12	Q9XWP9	Q9xwp9 coltivirus
151	6	3.8	181	7	Q99693	Q99693 homo sapien	224	6	3.8	264	2	Q9XA70	Q9xa70 streptomyce
152	6	3.8	181	7	Q99695	Q99695 homo sapien	225	6	3.8	264	7	Q29970	Q29970 homo sapien
153	6	3.8	181	7	Q99696	Q99696 homo sapien	226	6	3.8	265	5	Q27060	Q27060 trypanosoma
154	6	3.8	186	2	Q9RWM3	Q9rwm3 deinococcus	227	6	3.8	265	9	Q38155	Q38155 bacterioph
155	6	3.8	187	4	Q9BQD5	Q9bqd5 zymomonas m	228	6	3.8	266	2	Q9A6S6	Q9a6s6 caulobacter
156	6	3.8	187	4	Q9BQD5	Q9bqd5 homo sapien	229	6	3.8	267	2	Q9L044	Q9l044 streptomyce
157	6	3.8	188	2	Q83722	Q83722 treponema p	230	6	3.8	267	9	Q99363	Q99363 bacterioph
158	6	3.8	189	1	Q58399	Q58399 pyrococcus	231	6	3.8	268	5	Q45768	Q45768 caenorhabdi
159	6	3.8	190	2	P71658	P71658 mycobacteri	232	6	3.8	269	2	Q82875	Q82875 streptococc
160	6	3.8	191	3	Q59705	Q59705 schizosacch	233	6	3.8	269	7	Q30155	Q30155 homo sapien
161	6	3.8	191	4	Q9UEL2	Q9uel2 homo sapien	234	6	3.8	272	2	P96452	P96452 rhizobium m
162	6	3.8	193	10	Q9M163	Q9m163 arabidopsis	235	6	3.8	272	5	Q9VGA5	Q9vga5 pseudomonas
163	6	3.8	195	10	Q65256	Q65256 arabidopsis	236	6	3.8	272	5	Q9VGA5	Q9vga5 drosophila
164	6	3.8	196	2	Q9CNV9	Q9cnv9 pasteurella	237	6	3.8	273	4	Q9VWV3	Q9vww3 mycobacteri
165	6	3.8	196	13	Q12980	Q12980 fugu rubrip	238	6	3.8	273	4	Q06933	Q06933 homo sapien
			197	2	Q9RC58	Q9rcs8 synechococ							Q9hb16 homo sapien

239	6	3.8	273	7	077956	077956 homo sapien	312	6	3.8	278	5	094947	094947 drosophila
240	6	3.8	273	7	078150	078150 homo sapien	313	6	3.8	279	2	0990T6	0990T6 staphylococ
241	6	3.8	273	7	078151	078151 homo sapien	314	6	3.8	281	2	09A381	09A381 caulobacter
242	6	3.8	273	7	078152	078152 homo sapien	315	6	3.8	282	5	062174	062174 caenorhabdi
243	6	3.8	273	7	078153	078153 homo sapien	316	6	3.8	283	2	09RR10	09RR10 deinococcus
244	6	3.8	273	7	097987	097987 homo sapien	317	6	3.8	283	5	015928	015928 theileria o
245	6	3.8	273	7	097989	097989 homo sapien	318	6	3.8	283	5	076175	076175 theileria b
246	6	3.8	273	7	097990	097990 homo sapien	319	6	3.8	283	5	045767	045767 caenorhabdi
247	6	3.8	273	7	097991	097991 homo sapien	320	6	3.8	283	5	026792	026792 theileria b
248	6	3.8	273	7	097992	097992 homo sapien	321	6	3.8	283	5	09NDX5	09NDX5 theileria b
249	6	3.8	273	7	097993	097993 homo sapien	322	6	3.8	283	5	09NDV8	09NDV8 theileria b
250	6	3.8	273	7	097994	097994 homo sapien	323	6	3.8	283	5	09XWC9	09XWC9 caenorhabdi
251	6	3.8	273	7	097995	097995 homo sapien	324	6	3.8	284	11	09CSZ8	09CSZ8 mus musculu
252	6	3.8	273	7	098234	098234 homo sapien	325	6	3.8	286	2	09HV33	09HV33 pseudomonas
253	6	3.8	273	7	098235	098235 homo sapien	326	6	3.8	288	2	051082	051082 nocardia la
254	6	3.8	273	7	098236	098236 homo sapien	327	6	3.8	288	13	0919W8	0919W8 dendragapus
255	6	3.8	273	7	098238	098238 homo sapien	328	6	3.8	290	2	050903	050903 borrella bu
256	6	3.8	273	7	09TQL4	09TQL4 homo sapien	329	6	3.8	291	1	09HMD8	09HMD8 halobacteri
257	6	3.8	273	7	09TQL3	09TQL3 homo sapien	330	6	3.8	291	2	09JPD4	09JPD4 rhodocyclu
258	6	3.8	273	7	09TQL2	09TQL2 homo sapien	331	6	3.8	293	11	09D140	09D140 mus musculu
259	6	3.8	273	7	09TQL1	09TQL1 homo sapien	332	6	3.8	295	3	094321	094321 schizosacch
260	6	3.8	273	7	09TQL0	09TQL0 homo sapien	333	6	3.8	295	11	09N311	09N311 mus musculu
261	6	3.8	273	7	09TQK9	09TQK9 homo sapien	334	6	3.8	297	1	09HIK9	09HIK9 thermoplasm
262	6	3.8	273	7	09TQJ1	09TQJ1 homo sapien	335	6	3.8	298	2	09RJS2	09RJS2 streptomyce
263	6	3.8	273	7	09TQJ0	09TQJ0 homo sapien	336	6	3.8	298	5	09NH06	09NH06 heliothis z
264	6	3.8	273	7	09TQJ9	09TQJ9 homo sapien	337	6	3.8	301	2	049942	049942 mycobacteri
265	6	3.8	273	7	09TQI9	09TQI9 homo sapien	338	6	3.8	303	10	09C727	09C727 arabidopsis
266	6	3.8	273	7	09TQ67	09TQ67 homo sapien	339	6	3.8	304	2	09RRU2	09RRU2 deinococcus
267	6	3.8	274	4	060917	060917 homo sapien	340	6	3.8	305	2	09CNQ4	09CNQ4 pasteurella
268	6	3.8	274	4	060918	060918 homo sapien	341	6	3.8	305	10	09MIC3	09MIC3 arabidopsis
269	6	3.8	274	4	060919	060919 homo sapien	342	6	3.8	306	5	020135	020135 caenorhabdi
270	6	3.8	274	4	060920	060920 homo sapien	343	6	3.8	306	11	09D0Q7	09D0Q7 mus musculu
271	6	3.8	274	4	09UE11	09UE11 homo sapien	344	6	3.8	307	2	045418	045418 burkholderi
272	6	3.8	274	4	09UE06	09UE06 homo sapien	345	6	3.8	308	7	09TQK0	09TQK0 homo sapien
273	6	3.8	274	4	09UE05	09UE05 homo sapien	346	6	3.8	308	7	09TQK3	09TQK3 homo sapien
274	6	3.8	274	4	09UE04	09UE04 homo sapien	347	6	3.8	308	7	09TQ37	09TQ37 homo sapien
275	6	3.8	274	4	09UE03	09UE03 homo sapien	348	6	3.8	308	7	09BCS2	09BCS2 homo sapien
276	6	3.8	274	4	09UE01	09UE01 homo sapien	349	6	3.8	309	2	09RNN7	09RNN7 zymomonas m
277	6	3.8	274	4	09UE00	09UE00 homo sapien	350	6	3.8	310	2	09PMF4	09PMF4 campylobact
278	6	3.8	274	4	09UD29	09UD29 homo sapien	351	6	3.8	312	4	09GZK6	09GZK6 homo sapien
279	6	3.8	274	4	09UD28	09UD28 homo sapien	352	6	3.8	312	4	09GZK1	09GZK1 homo sapien
280	6	3.8	274	7	079506	079506 homo sapien	353	6	3.8	316	1	09HM43	09HM43 thermoplasm
281	6	3.8	274	7	079507	079507 homo sapien	354	6	3.8	317	7	09TQJ9	09TQJ9 homo sapien
282	6	3.8	274	7	079510	079510 homo sapien	355	6	3.8	317	7	09TQ95	09TQ95 homo sapien
283	6	3.8	274	7	079512	079512 homo sapien	356	6	3.8	317	7	09TQ92	09TQ92 homo sapien
284	6	3.8	274	7	079516	079516 homo sapien	357	6	3.8	317	7	09G1Z3	09G1Z3 homo sapien
285	6	3.8	274	7	079517	079517 homo sapien	358	6	3.8	318	1	09HSH7	09HSH7 halobacteri
286	6	3.8	274	7	079518	079518 homo sapien	359	6	3.8	318	7	09TQK2	09TQK2 homo sapien
287	6	3.8	274	7	079520	079520 homo sapien	360	6	3.8	318	7	09TQK1	09TQK1 homo sapien
288	6	3.8	274	7	099692	099692 homo sapien	361	6	3.8	318	7	09TQK4	09TQK4 homo sapien
289	6	3.8	274	7	079521	079521 homo sapien	362	6	3.8	318	7	09TQK0	09TQK0 homo sapien
290	6	3.8	274	7	079515	079515 homo sapien	363	6	3.8	318	7	09TQ93	09TQ93 homo sapien
291	6	3.8	274	7	079514	079514 homo sapien	364	6	3.8	318	7	09GJ03	09GJ03 homo sapien
292	6	3.8	274	7	079513	079513 homo sapien	365	6	3.8	318	7	09G1Z6	09G1Z6 homo sapien
293	6	3.8	274	7	09TQ89	09TQ89 homo sapien	366	6	3.8	318	7	09G1Z4	09G1Z4 homo sapien
294	6	3.8	274	7	09TQ88	09TQ88 homo sapien	367	6	3.8	319	5	09NE45	09NE45 leishmania
295	6	3.8	274	7	09TQ87	09TQ87 homo sapien	368	6	3.8	319	5	09N538	09N538 caenorhabdi
296	6	3.8	274	7	09TQ86	09TQ86 homo sapien	369	6	3.8	319	7	09TQK2	09TQK2 homo sapien
297	6	3.8	274	12	09IH19	09IH19 grass carp	370	6	3.8	319	7	09TQ98	09TQ98 homo sapien
298	6	3.8	275	2	050711	050711 mycobacteri	371	6	3.8	319	7	09TQ94	09TQ94 homo sapien
299	6	3.8	275	2	09F200	09F200 bacillus ci	372	6	3.8	321	10	09SZF6	09SZF6 arabidopsis
300	6	3.8	275	2	099VF4	099VF4 staphylococ	373	6	3.8	322	7	09TQ96	09TQ96 homo sapien
301	6	3.8	275	7	098019	098019 gorilla gor	374	6	3.8	322	7	09G1Z5	09G1Z5 homo sapien
302	6	3.8	275	7	098020	098020 pan paniscu	375	6	3.8	322	7	09BD20	09BD20 homo sapien
303	6	3.8	275	7	098021	098021 pongo pygma	376	6	3.8	323	2	09HZT4	09HZT4 pseudomonas
304	6	3.8	275	7	098022	098022 pongo pygma	377	6	3.8	323	5	09VMG4	09VMG4 drosophila
305	6	3.8	277	2	032904	032904 mycobacteri	378	6	3.8	323	7	09MY21	09MY21 homo sapien
306	6	3.8	277	2	09K420	09K420 streptomyce	379	6	3.8	323	10	09STS2	09STS2 arabidopsis
307	6	3.8	278	5	094852	094852 drosophila	380	6	3.8	327	2	P76054	P76054 escherichia
308	6	3.8	278	5	094894	094894 drosophila	381	6	3.8	327	2	09EYX5	09EYX5 salmonella
309	6	3.8	278	5	094921	094921 drosophila	382	6	3.8	328	2	09CE05	09CE05 lactococcus
310	6	3.8	278	5	094925	094925 drosophila	383	6	3.8	329	2	045987	045987 caulobacter
311	6	3.8	278	5	094926	094926 drosophila	384	6	3.8	329	5	Q17166	Q17166 brugia mala

385	6	3.8	330	10	Q9FV74	Q9fvt4 arabidopsis	458	6	3.8	401	10	Q9ZPI4	Q9zpi4 nicotiana p
386	6	3.8	331	10	O04200	O04200 arabidopsis	459	6	3.8	402	2	Q9WZK6	Q9wzk6 thermotoga
387	6	3.8	331	11	Q9YL63	Q9yl63 mus musculus	460	6	3.8	403	4	Q9H620	Q9h620 homo sapien
388	6	3.8	333	2	Q9KWD6	Q9kwd6 agrobacteri	461	6	3.8	404	2	Q9PMY9	Q9pm9 campylobact
389	6	3.8	333	4	Q9BV77	Q9bv77 homo sapien	462	6	3.8	404	2	Q9PMW4	Q9pm4 campylobact
390	6	3.8	333	10	Q9X123	Q9x123 oryza sativ	463	6	3.8	405	2	P74479	P74479 synechocyst
391	6	3.8	336	2	Q9RDM2	Q9rdm2 streptomyce	464	6	3.8	406	2	P96564	P96564 amycolatops
392	6	3.8	336	2	Q9F2Y8	Q9f2y8 streptomyce	465	6	3.8	408	2	Q9PMW3	Q9pmw3 campylobact
393	6	3.8	338	2	Q9S699	Q9s699 streptomyce	466	6	3.8	408	2	Q9PGZ7	Q9pgz7 xylella fas
394	6	3.8	338	2	Q9EWS3	Q9ews3 streptomyce	467	6	3.8	409	2	Q9PGZ7	Q9pgz7 xylella fas
395	6	3.8	339	2	Q9JN80	Q9jnn80 streptomyce	468	6	3.8	412	5	Q26806	Q26806 trypanosoma
396	6	3.8	340	2	Q9I2F0	Q9i2f0 pseudomonas	469	6	3.8	412	11	Q9CU83	Q9cu83 mus musculus
397	6	3.8	341	2	Q9XAM1	Q9xam1 streptomyce	470	6	3.8	414	2	Q9Z7D0	Q9z7d0 chlamydia p
398	6	3.8	341	2	Q9HVG0	Q9hvg0 pseudomonas	471	6	3.8	415	2	Q94105	Q94105 mycobacteri
399	6	3.8	341	2	Q9A443	Q9a443 caulobacter	472	6	3.8	415	5	Q76907	Q76907 drosophila
400	6	3.8	342	12	Q9DWD1	Q9dwd1 rat cytomeg	473	6	3.8	415	5	Q9W4V8	Q9w4v8 drosophila
401	6	3.8	343	2	Q9RRE0	Q9rre0 delnococcus	474	6	3.8	415	5	Q9U4M9	Q9u4m9 leishmania
402	6	3.8	344	10	Q9ARW2	Q9arw2 oryza sativ	475	6	3.8	416	2	Q87605	Q87605 streptomyce
403	6	3.8	345	5	P91014	P91014 caenorhabdi	476	6	3.8	419	4	Q9H1L8	Q9h1l8 homo sapien
404	6	3.8	345	6	Q27951	Q27951 bos taurus	477	6	3.8	420	2	Q9X845	Q9x845 streptomyce
405	6	3.8	351	2	Q9RKP3	Q9rkp3 streptomyce	478	6	3.8	421	2	Q9JYQ5	Q9jyq5 neisseria m
406	6	3.8	351	2	O52572	O52572 amycolatops	479	6	3.8	421	2	Q9JTP6	Q9jtp6 neisseria m
407	6	3.8	351	5	O45256	O45256 caenorhabdi	480	6	3.8	422	2	Q9L011	Q9l011 streptomyce
408	6	3.8	352	2	Q9AMU2	Q9amu2 bradyrhizob	481	6	3.8	423	10	Q43450	Q43450 glycine max
409	6	3.8	353	2	Q9PCQ1	Q9pcq1 xylella fas	482	6	3.8	425	2	Q9RNN4	Q9rnn4 zymonocus m
410	6	3.8	354	2	Q9I5Y7	Q9i5y7 pseudomonas	483	6	3.8	426	2	Q9RS78	Q9rs78 delnococcus
411	6	3.8	355	1	Q9UXH7	Q9uxh7 sulfolobus	484	6	3.8	428	5	Q24024	Q24024 drosophila
412	6	3.8	358	2	Q9I1U3	Q9i1u3 pseudomonas	485	6	3.8	428	5	Q9VPT3	Q9vpt3 drosophila
413	6	3.8	359	7	Q9BCR7	Q9bcr7 homo sapien	486	6	3.8	428	10	O81641	O81641 arabidopsis
414	6	3.8	359	7	Q9BCR5	Q9bcr5 homo sapien	487	6	3.8	433	2	O85304	O85304 salmonella
415	6	3.8	359	7	Q9BCR3	Q9bcr3 homo sapien	488	6	3.8	435	13	Q9I9W9	Q9i9w9 dendragapus
416	6	3.8	360	3	Q05791	Q05791 saccharomyc	489	6	3.8	441	2	O85201	O85201 streptococc
417	6	3.8	360	7	Q9BCS0	Q9bcs0 homo sapien	490	6	3.8	443	2	Q9JZC9	Q9jzc9 neisseria m
418	6	3.8	360	7	Q9BCR8	Q9bcr8 homo sapien	491	6	3.8	444	13	Q9VHC6	Q9vhc6 rana ridibu
419	6	3.8	360	7	Q9BCR6	Q9bcr6 homo sapien	492	6	3.8	445	2	Q9RD63	Q9rd63 bacillus ha
420	6	3.8	361	1	Q9YFD2	Q9yfd2 aeropyrum p	493	6	3.8	448	5	P91064	P91064 caenorhabdi
421	6	3.8	361	7	Q9BCS1	Q9bcs1 homo sapien	494	6	3.8	451	10	Q9SB50	Q9sb50 arabidopsis
422	6	3.8	361	7	Q9BCR9	Q9bcr9 homo sapien	495	6	3.8	451	10	Q9WF08	Q9wf08 streptomyce
423	6	3.8	362	2	Q9A7D6	Q9a7d6 caulobacter	496	6	3.8	461	10	Q9FUI7	Q9fui7 oryza sativ
424	6	3.8	363	2	Q9Z514	Q9z514 streptomyce	497	6	3.8	461	10	Q9FUI7	Q9fui7 oryza sativ
425	6	3.8	363	2	Q9NY58	Q9ny58 homo sapien	498	6	3.8	461	10	Q9FUI7	Q9fui7 oryza sativ
426	6	3.8	364	7	Q9BCS4	Q9bcs4 homo sapien	500	6	3.8	461	10	Q9FUI7	Q9fui7 oryza sativ
427	6	3.8	364	7	Q9BCR4	Q9bcr4 homo sapien	501	6	3.8	461	10	Q9FUI7	Q9fui7 oryza sativ
428	6	3.8	365	4	Q9H1L9	Q9h1l9 homo sapien	502	6	3.8	461	10	Q9FUI7	Q9fui7 oryza sativ
429	6	3.8	365	10	Q43451	Q43451 glycine max	503	6	3.8	461	10	Q9FUI7	Q9fui7 oryza sativ
430	6	3.8	366	6	Q9NLP5	Q9nlp5 bubalus bub	504	6	3.8	463	5	Q19511	Q19511 caenorhabdi
431	6	3.8	369	2	Q9PGQ4	Q9pgq4 xylella fas	505	6	3.8	464	1	Q9YA63	Q9ya63 aeropyrum p
432	6	3.8	371	2	Q9HZW5	Q9hzw5 pseudomonas	506	6	3.8	464	1	Q9YA63	Q9ya63 aeropyrum p
433	6	3.8	371	4	Q9NRP3	Q9nrp3 homo sapien	507	6	3.8	466	5	O17394	O17394 caenorhabdi
434	6	3.8	374	5	Q9VJA0	Q9vja0 drosophila	508	6	3.8	468	4	O00220	O00220 homo sapien
435	6	3.8	375	2	Q9FCC4	Q9fcc4 streptomyce	509	6	3.8	469	2	Q9KN29	Q9kn29 vibrio chol
436	6	3.8	375	7	Q98265	Q98265 pan troglod	510	6	3.8	470	10	Q9FFF1	Q9fff1 arabidopsis
437	6	3.8	376	4	Q15635	Q15635 homo sapien	511	6	3.8	471	5	Q9V7R8	Q9v7r8 drosophila
438	6	3.8	376	10	Q9LZK8	Q9lzk8 arabidopsis	512	6	3.8	472	2	Q9S6J9	Q9s6j9 streptomyce
439	6	3.8	377	1	O30022	O30022 archaeoglob	513	6	3.8	473	2	Q9A3N2	Q9a3n2 caulobacter
440	6	3.8	377	5	Q9NB18	Q9nb18 drosophila	514	6	3.8	473	2	Q9A3N2	Q9a3n2 caulobacter
441	6	3.8	378	2	O86863	O86863 streptomyce	515	6	3.8	475	10	O81325	O81325 arabidopsis
442	6	3.8	378	6	Q9TTY4	Q9ttv4 canis famil	516	6	3.8	477	5	Q9VZT0	Q9vzt0 drosophila
443	6	3.8	380	10	O48535	O48535 arabidopsis	517	6	3.8	480	2	Q9JW97	Q9jw97 neisseria m
444	6	3.8	383	2	Q9KDX4	Q9kdx4 bacillus ha	518	6	3.8	482	2	Q69046	Q69046 rhodococcus
445	6	3.8	383	5	P91439	P91439 caenorhabdi	519	6	3.8	482	2	Q69046	Q69046 rhodococcus
446	6	3.8	383	7	Q29983	Q29983 homo sapien	520	6	3.8	486	4	Q9HA78	Q9ha78 homo sapien
447	6	3.8	385	2	Q9X4F1	Q9x4f1 rhodobacter	521	6	3.8	489	11	Q9WUH5	Q9wuh5 mus musculu
448	6	3.8	387	2	O69208	O69208 actinosynne	522	6	3.8	489	11	Q9CY03	Q9cy03 mus musculu
449	6	3.8	387	2	O56076	O56076 streptomyce	523	6	3.8	492	10	O80830	O80830 arabidopsis
450	6	3.8	388	2	Q9KE27	Q9ke27 bacillus ha	524	6	3.8	493	2	Q9EV50	Q9ev50 staphylococ
451	6	3.8	389	2	Q9AN18	Q9an18 bradyrhizob	525	6	3.8	493	5	O17554	O17554 caenorhabdi
452	6	3.8	389	10	Q9FEW4	Q9few4 medicago tr	526	6	3.8	494	11	Q17554	Q17554 caenorhabdi
453	6	3.8	390	4	Q9UJG8	Q9ujg8 homo sapien	527	6	3.8	494	11	Q17554	Q17554 caenorhabdi
454	6	3.8	396	1	O28239	O28239 archaeoglob	528	6	3.8	496	2	O05896	O05896 mycobacteri
455	6	3.8	398	2	Q9RR92	Q9rr92 delnococcus	529	6	3.8	499	5	Q9GZ34	Q9gz34 trypanosoma
456	6	3.8	398	13	Q9I9U1	Q9i9u1 lagopus leu	530	6	3.8	502	2	O53551	O53551 mycobacteri
457	6	3.8	400	2	Q48814	Q48814 legionella				503	12	Q9WNM4	Q9wnm4 human papil

531	6	3.8	504	10	O80515	O80515 arabidopsis	604	2	Q9S562	Q9S562 pseudomonas
532	6	3.8	504	10	Q9LDF0	Q9Ldf0 oryza sativ	605	4	Q9NTN3	Q9ntn3 homo sapien
533	6	3.8	505	2	Q9KW19	Q9kw19 staphylococ	606	5	Q9Y095	Q9y095 drosophila
534	6	3.8	506	4	Q9H694	Q9h694 homo sapien	607	5	Q9UB07	Q9ub07 drosophila
535	6	3.8	507	1	Q26749	O26749 methanobact	608	2	Q49182	Q49182 mycobacteri
536	6	3.8	508	10	Q9LDQ4	Q9ldq4 oryza sativ	609	6	Q67340	Q67340 aquifex aeo
537	6	3.8	510	1	Q9Y9C5	Q9y9c5 aeropyrum p	610	6	Q9P938	Q9p938 pichia stip
538	6	3.8	510	2	Q9F1M3	Q9flm3 synechococc	611	6	Q94263	Q94263 caenorhabdi
539	6	3.8	511	2	Q9X5T0	Q9x5t0 streptomyce	612	6	Q58099	Q58099 pyrococcus
540	6	3.8	511	10	Q9LWX3	Q9lwx3 oryza sativ	613	5	Q9VK21	Q9vk21 drosophila
541	6	3.8	511	10	Q9LE72	Q9le72 oryza sativ	614	5	Q9V877	Q9v877 drosophila
542	6	3.8	512	2	Q9JZ91	Q9jz91 neisseria m	615	6	Q9L8G0	Q9l8g0 streptomyce
543	6	3.8	512	2	Q9JUB5	Q9jub5 neisseria m	616	6	Q9L8G0	Q9l8g0 streptomyce
544	6	3.8	512	10	Q9LDC7	Q9ldc7 oryza sativ	617	6	Q9CDA6	Q9cda6 mycobacteri
545	6	3.8	513	3	Q9C1X4	Q9c1x4 schizosacch	618	6	Q9LIQ7	Q9liq7 arabidopsis
546	6	3.8	515	3	Q00056	Q00056 penicillium	619	6	Q9NSR6	Q9nsr6 homo sapien
547	6	3.8	515	3	Q9HFA5	Q9hfa5 penicillium	620	6	Q9Z389	Q9z389 streptomyce
548	6	3.8	516	3	Q74641	O74641 aspergillus	621	6	Q9Z1F9	Q9z1f9 mus musculu
549	6	3.8	516	3	Q74642	O74642 aspergillus	622	6	Q9UBT2	Q9ubt2 homo sapien
550	6	3.8	516	3	Q94220	O94220 aspergillus	623	6	Q9NTJ1	Q9ntj1 homo sapien
551	6	3.8	516	4	Q9H647	Q9h647 homo sapien	624	6	Q95605	Q95605 homo sapien
552	6	3.8	517	2	Q9S2Q7	Q9s2q7 streptomyce	625	6	Q67389	Q67389 aquifex aeo
553	6	3.8	520	5	Q9VLC5	Q9vlc5 drosophila	626	6	Q9F2W6	Q9f2w6 streptomyce
554	6	3.8	521	13	P79923	P79923 xenopus lae	627	6	Q9PAN5	Q9pan5 xylella fas
555	6	3.8	523	10	Q9LV40	Q9lv40 arabidopsis	628	6	Q9KQK8	Q9kqk8 vibrio chol
556	6	3.8	525	13	Q9PVH8	Q9pvh8 xenopus lae	629	6	Q9CALK3	Q9cal3 arabidopsis
557	6	3.8	529	10	Q9FZ11	Q9fz11 arachis hyp	630	6	Q9UF93	Q9uf93 homo sapien
558	6	3.8	530	10	Q9SOH7	Q9soh7 arachis hyp	631	6	Q9ZUT5	Q9zut5 arabidopsis
559	6	3.8	534	5	O76843	O76843 caenorhabdi	632	6	Q9ZUH6	Q9zuh6 arabidopsis
560	6	3.8	540	5	Q9VFS7	Q9vfs7 drosophila	633	6	Q9TUQ1	Q9tuq1 bos taurus
561	6	3.8	540	9	O21870	O21870 bacterioph	634	6	Q9UVV2	Q9uvv2 aspergillus
562	6	3.8	540	9	O80104	O80104 bacterioph	635	6	Q9N8P1	Q9n8p1 trypanosoma
563	6	3.8	540	9	O80279	O80279 lactococcus	636	6	Q15211	Q15211 homo sapien
564	6	3.8	541	2	Q9K2X5	Q9k2x5 streptomyce	637	6	Q9H950	Q9h950 homo sapien
565	6	3.8	544	5	Q20780	Q20780 caenorhabdi	638	6	Q9HCU8	Q9hcj8 homo sapien
566	6	3.8	545	5	O18276	O18276 caenorhabdi	639	6	Q9V7V6	Q9v7v6 drosophila
567	6	3.8	545	11	O88836	O88836 mus musculu	640	10	Q9LW73	Q9lw73 arabidopsis
568	6	3.8	548	10	Q9SVE2	Q9sve2 arabidopsis	641	6	Q92615	Q92615 homo sapien
569	6	3.8	550	2	Q9S2A6	Q9s2a6 streptomyce	642	5	Q23706	Q23706 carcinus ma
570	6	3.8	553	4	Q13517	Q13517 homo sapien	643	2	O33614	O33614 streptomyce
571	6	3.8	556	5	Q9VLB8	Q9vlb8 drosophila	644	2	Q9ZBK2	Q9zbk2 streptomyce
572	6	3.8	556	10	Q9FFK3	Q9ffk3 arabidopsis	645	2	Q9EZ60	Q9ez60 streptomyce
573	6	3.8	559	10	Q9FMW3	Q9fmw3 arabidopsis	646	6	Q9V7V7	Q9v7v7 drosophila
574	6	3.8	560	2	Q9CHK0	Q9chk0 lactococcus	647	5	O62527	O62527 drosophila
575	6	3.8	562	2	O05305	O05305 mycobacteri	648	5	Q9TYZ7	Q9tyz7 caenorhabdi
576	6	3.8	565	10	O22511	O22511 vitis vinif	649	2	Q9KJ07	Q9kj07 mycobacteri
577	6	3.8	566	2	O86854	O86854 streptomyce	650	6	Q9P215	Q9p215 homo sapien
578	6	3.8	566	5	Q9VYE8	Q9vye8 drosophila	651	6	Q9V5G3	Q9v5g3 drosophila
579	6	3.8	566	12	O36454	O36454 lymantria d	652	6	Q9UH90	Q9uh90 homo sapien
580	6	3.8	568	2	Q9ZHI2	Q9zhi2 chromobacte	653	6	Q24205	Q24205 drosophila
581	6	3.8	569	3	Q9HEE7	Q9hee7 neurospora	654	2	Q9X5N7	Q9x5n7 escherichia
582	6	3.8	569	10	Q9FJQ6	Q9fjq6 arabidopsis	655	10	O04721	O04721 arabidopsis
583	6	3.8	571	10	Q9F192	Q9fi92 arabidopsis	656	6	Q9ULM5	Q9ulm5 homo sapien
584	6	3.8	573	1	Q9P9D5	Q9p9d5 uncultured	657	3	Q9UTQ2	Q9utq2 schizosacch
585	6	3.8	576	2	O06137	O06137 mycobacteri	658	719	Q9HFJ6	Q9hjf6 neurospora
586	6	3.8	578	3	Q9UT49	Q9ut49 schizosacch	659	5	Q9U9U5	Q9u9u5 dictyosteli
587	6	3.8	579	10	Q9M056	Q9m056 arabidopsis	660	12	P89466	P89466 herpes simp
588	6	3.8	581	1	Q9P9C2	Q9p9c2 uncultured	661	10	Q9M075	Q9m075 arabidopsis
589	6	3.8	583	4	O15290	O15290 homo sapien	662	10	Q9M075	Q9m075 arabidopsis
590	6	3.8	583	10	P93264	P93264 mesembryant	663	12	Q00279	Q00279 foot and mo
591	6	3.8	585	12	O89244	O89244 woodchuck h	664	5	Q9V749	Q9vt49 drosophila
592	6	3.8	586	5	Q22551	Q22551 caenorhabdi	665	5	Q46076	Q46076 drosophila
593	6	3.8	587	10	Q38954	Q38954 arabidopsis	666	5	Q9V4J5	Q9v4j5 drosophila
594	6	3.8	588	10	Q9FT12	Q9ft12 brassica na	667	2	Q9X7M0	Q9x7m0 rhodobacter
595	6	3.8	589	10	Q43390	Q43390 brassica na	668	5	Q9VVZ0	Q9vvz0 drosophila
596	6	3.8	591	11	P70590	P70590 rattus norv	669	4	Q9H2G5	Q9h2g5 homo sapien
597	6	3.8	592	2	Q9I610	Q9i610 pseudomonas	670	1	Q9U2F5	Q9uzf5 pyrococcus
598	6	3.8	594	5	Q9BLH5	Q9blh5 halocynthia	671	6	O86766	O86766 streptomyce
599	6	3.8	596	4	O15464	O15464 homo sapien	672	2	O30329	O30329 acetobacter
600	6	3.8	599	5	Q22089	Q22089 caenorhabdi	673	3	Q9HEC5	Q9hec5 neurospora
601	6	3.8	604	2	Q9K3R7	Q9k3r7 streptomyce	674	4	Q15634	Q15634 homo sapien
602	6	3.8	607	10	Q9SI43	Q9si43 arabidopsis	675	5	Q9VBS6	Q9vbs6 drosophila
603	6	3.8	609	10	Q9LR69	Q9lr69 arabidopsis	676	6	Q9XSX1	Q9xsx1 bos taurus

677	6	3.8	797	2	Q9RKS9	Q9rks9 streptomyc	750	6	3.8	1098	4	Q9UPQ7	Q9upq7 homo sapien
678	6	3.8	798	11	O88568	O88568 mus musculus	751	6	3.8	1107	11	Q63356	Q63356 rattus norv
679	6	3.8	798	11	O63555	Q63555 rattus norv	752	6	3.8	1109	4	Q12965	Q12965 homo sapien
680	6	3.8	799	10	Q9LUS1	Q9lus1 arabidopsis	753	6	3.8	1113	4	Q9ULL1	Q9ull1 homo sapien
681	6	3.8	802	4	Q9H998	Q9h998 homo sapien	754	6	3.8	1131	12	Q65030	Q65030 arabis mosa
682	6	3.8	806	4	Q9BQ09	Q9bq09 homo sapien	755	6	3.8	1134	1	Q30078	Q30078 archaeoglob
683	6	3.8	808	4	Q9UR88	Q9uk88 homo sapien	756	6	3.8	1139	2	Q50466	Q50466 mycobacteri
684	6	3.8	813	3	Q9UQW1	Q9uqw1 saccharomyc	757	6	3.8	1157	1	Q9HH42	Q9hh42 methanobact
685	6	3.8	816	12	Q99GX2	Q99gx2 helicoverp	758	6	3.8	1179	4	Q14120	Q14120 homo sapien
686	6	3.8	822	10	Q9FNM8	Q9fnm8 arabidopsis	759	6	3.8	1179	5	Q9V450	Q9v450 drosophila
687	6	3.8	827	12	Q65382	Q65382 broad bean	760	6	3.8	1194	3	O42649	O42649 schizosacch
688	6	3.8	827	12	Q65383	Q65383 broad bean	761	6	3.8	1198	5	Q9U5Y2	Q9u5y2 drosophila
689	6	3.8	839	5	Q9Y103	Q9y103 drosophila	762	6	3.8	1200	5	Q9W8U8	Q9w8u8 drosophila
690	6	3.8	839	5	Q9NHN7	Q9nhn7 drosophila	763	6	3.8	1200	5	Q9U531	Q9u531 drosophila
691	6	3.8	835	6	Q9TUQ0	Q9tuq0 bos taurus	764	6	3.8	1200	5	Q9U472	Q9u472 drosophila
692	6	3.8	857	10	Q9SBZ0	Q9sbz0 phaseolus a	765	6	3.8	1201	4	Q9P2D7	Q9p2d7 homo sapien
693	6	3.8	858	12	Q9ELI5	Q9eli5 meleagrid h	766	6	3.8	1205	4	Q9ULU4	Q9ulu4 homo sapien
694	6	3.8	858	12	Q9DPS9	Q9dps9 meleagrid h	767	6	3.8	1207	4	Q9C095	Q9c095 homo sapien
695	6	3.8	863	8	Q9TL05	Q9tli05 nephroselmi	768	6	3.8	1214	5	Q9VUX2	Q9vux2 drosophila
696	6	3.8	867	12	Q9WQ04	Q9wg04 simian retr	769	6	3.8	1215	10	Q9S151	Q9s151 arabidopsis
697	6	3.8	867	12	Q9WQ00	Q9wg00 simian retr	770	6	3.8	1219	10	Q9AY59	Q9ay59 oryza sativ
698	6	3.8	870	4	Q9NSA5	Q9nsa5 homo sapien	771	6	3.8	1234	11	Q9R044	Q9r044 rattus norv
699	6	3.8	870	10	O81293	O81293 arabidopsis	772	6	3.8	1235	4	O95428	O95428 homo sapien
700	6	3.8	886	2	Q9KC72	Q9kc72 bacillus ha	773	6	3.8	1237	12	O56042	O56042 rice tagged
701	6	3.8	887	2	Q9CMD5	Q9cmd5 pasteurella	774	6	3.8	1240	12	Q9DWH8	Q9dwh8 rat cytomeg
702	6	3.8	889	10	Q9FK48	Q9fk48 arabidopsis	775	6	3.8	1242	11	Q9QZS7	Q9qzs7 mus musculus
703	6	3.8	891	5	O24926	O24926 eupryma sc	776	6	3.8	1243	2	Q9RDG5	Q9rdg5 streptomyc
704	6	3.8	891	5	Q9GU16	Q9gu16 caenorhabdi	777	6	3.8	1245	3	Q9Y7V5	Q9y7v5 trichoderma
705	6	3.8	894	3	Q93833	Q93833 saccharomyc	778	6	3.8	1247	5	Q9YOT2	Q9yot2 dictyostell
706	6	3.8	894	5	O24925	O24925 eupryma sc	779	6	3.8	1247	5	Q9U9T1	Q9u9t1 dictyostell
707	6	3.8	906	11	O55129	O55129 mus musculus	780	6	3.8	1252	11	Q9QXX7	Q9qxx7 rattus norv
708	6	3.8	910	4	Q9Y620	Q9y620 homo sapien	781	6	3.8	1252	11	Q9JIX2	Q9jix2 rattus norv
709	6	3.8	910	12	Q9O915	Q9o915 avian adeno	782	6	3.8	1256	11	Q9JIX1	Q9jix1 mus musculus
710	6	3.8	916	3	Q94144	Q94144 saccharomyc	783	6	3.8	1256	11	O9PT59	O9pt59 mus musculus
711	6	3.8	917	4	Q9PIY5	Q9piy5 homo sapien	784	6	3.8	1274	5	O20007	O20007 caenorhabdi
712	6	3.8	926	5	O18504	O18504 septa offic	785	6	3.8	1283	5	O97034	O97034 ephydatia f
713	6	3.8	930	6	Q9XSW5	Q9xsw5 bos taurus	786	6	3.8	1326	5	O16928	O16928 caenorhabdi
714	6	3.8	945	2	Q9RD88	Q9rd88 streptomyc	787	6	3.8	1331	10	O49549	O49549 arabidopsis
715	6	3.8	947	5	O98667	O98667 plasmodium	788	6	3.8	1341	5	Q9GRK2	Q9grk2 leishmania
716	6	3.8	950	2	O50470	O50470 mycobacteri	789	6	3.8	1354	5	Q9VKA7	Q9vka7 drosophila
717	6	3.8	952	11	O63560	O63560 rattus norv	790	6	3.8	1378	10	O9LTA4	O9lta4 arabidopsis
718	6	3.8	956	11	O08794	O08794 mus musculus	791	6	3.8	1386	5	Q9GRS8	Q9grs8 leishmania
719	6	3.8	977	11	Q99MQ1	Q99mq1 mus musculus	792	6	3.8	1407	3	O06211	O06211 saccharomyc
720	6	3.8	978	2	P96104	P96104 thiobacilli	793	6	3.8	1420	3	Q06232	Q06232 trypanosoma
721	6	3.8	980	3	Q9P529	Q9p529 neurospora	794	6	3.8	1431	5	Q9NG75	Q9ng75 paratemia
722	6	3.8	983	2	Q9CCM6	Q9ccm6 mycobacteri	795	6	3.8	1433	2	P75020	P75020 streptomyc
723	6	3.8	992	2	O53339	O53339 mycobacteri	796	6	3.8	1433	2	Q9L040	Q9l040 streptomyc
724	6	3.8	1001	2	O25314	O25314 helicobacte	797	6	3.8	1445	10	O24438	O24438 oryza longi
725	6	3.8	1011	10	Q9STU0	Q9stu0 arabidopsis	798	6	3.8	1446	10	Q9SVT9	Q9svt9 arabidopsis
726	6	3.8	1011	11	Q9QY54	Q9qy54 mus musculus	799	6	3.8	1463	5	O9U5E2	O9u5e2 drosophila
727	6	3.8	1018	5	O17874	O17874 caenorhabdi	800	6	3.8	1464	11	O08948	O08948 rattus norv
728	6	3.8	1019	10	Q9LM78	Q9lm78 arabidopsis	801	6	3.8	1464	11	O63728	O63728 rattus norv
729	6	3.8	1021	5	Q9XUV2	Q9xuv2 caenorhabdi	802	6	3.8	1473	5	Q9BI87	Q9bi87 caenorhabdi
730	6	3.8	1022	5	Q9V6T7	Q9v6t7 drosophila	803	6	3.8	1479	11	O92277	O92277 mus musculus
731	6	3.8	1025	3	Q9UUX9	Q9uux9 neurospora	804	6	3.8	1502	10	O24517	O24517 helianthus
732	6	3.8	1025	3	O25693	O25693 plasmodium	805	6	3.8	1505	10	Q9LKB9	Q9lkb9 arabidopsis
733	6	3.8	1032	2	Q9CL57	Q9cls7 pasteurella	806	6	3.8	1515	10	Q39158	Q39158 arabidopsis
734	6	3.8	1033	2	O33373	O33373 neisseria g	807	6	3.8	1515	10	O9ATB7	O9atb7 petroselinu
735	6	3.8	1035	2	O33372	O33372 neisseria g	808	6	3.8	1516	10	Q9LH86	Q9lh86 arabidopsis
736	6	3.8	1037	2	P72083	P72083 neisseria m	809	6	3.8	1611	3	O42854	O42854 schizosacch
737	6	3.8	1037	2	Q9JY25	Q9jyv25 neisseria m	810	6	3.8	1655	5	O44498	O44498 caenorhabdi
738	6	3.8	1038	2	O05924	O05924 neisseria m	811	6	3.8	1685	4	Q9UVM8	Q9uem8 homo sapien
739	6	3.8	1038	3	Q9PAC6	Q9pac6 pichia past	812	6	3.8	1693	5	Q9VXV8	Q9vxv8 drosophila
740	6	3.8	1043	11	O99N32	O99n32 mus musculus	813	6	3.8	1733	2	O53579	O53579 mycobacteri
741	6	3.8	1044	2	O05923	O05923 neisseria g	814	6	3.8	1733	2	O53579	O53579 mycobacteri
742	6	3.8	1046	2	Q9ZNG8	Q9zng8 pseudomonas	815	6	3.8	1751	5	O17143	O17143 drosophila
743	6	3.8	1048	2	O05925	O05925 neisseria m	816	6	3.8	1770	12	O56224	O56224 simian maso
744	6	3.8	1050	2	O51019	O51019 neisseria g	817	6	3.8	1773	5	O17142	O17142 drosophila
745	6	3.8	1063	11	Q9OY55	Q9oy55 mus musculus	818	6	3.8	1783	10	Q9PEC4	Q9fec4 chlamydomon
746	6	3.8	1069	4	O9H6B5	O9h6b5 homo sapien	819	6	3.8	1784	3	Q9CDB1	Q9cdb1 mycobacteri
747	6	3.8	1074	10	Q9LLN7	Q9lln7 oryza sativ	820	6	3.8	1796	3	Q99313	Q99313 saccharomyc
748	6	3.8	1077	10	O81441	O81441 arabidopsis	821	6	3.8	1816	10	Q9SDB1	Q9sdb1 arabidopsis
749	6	3.8	1084	12	Q65029	Q65029 arabis mosa	822	6	3.8	1867	4	Q9UPY0	Q9upy0 homo sapien
										1875	5	Q93691	Q93691 caenorhabdi

823 6 3.8 1970 5 Q9VQ08  
824 6 3.8 1983 5 Q9UI65  
825 6 3.8 2088 5 Q9VXE6  
826 6 3.8 2115 12 Q86500  
827 6 3.8 2116 12 Q40955  
828 6 3.8 2116 12 Q991E7  
829 6 3.8 2116 12 Q991E5  
830 6 3.8 2165 5 Q9VH88  
831 6 3.8 2301 10 Q9ATK5  
832 6 3.8 2303 4 Q95996  
833 6 3.8 2340 5 Q97298  
834 6 3.8 2342 5 Q01677  
835 6 3.8 2442 4 Q60588  
836 6 3.8 2442 4 Q14812  
837 6 3.8 2442 4 Q9H450  
838 6 3.8 2523 2 Q53393  
839 6 3.8 2920 11 Q9R0M0  
840 6 3.8 3247 12 Q65553  
841 6 3.8 3455 2 Q9P906  
842 6 3.8 3583 2 Q45675  
843 6 3.8 3680 6 Q97592  
844 6 3.8 4688 2 Q9PQ08  
845 6 3.8 4904 5 Q9VH01  
846 6 3.8 5636 5 Q9N9M2

Q9vqu8 drosophila  
Q9ul65 leishmania  
Q9vxe6 drosophila  
Q86500 rubella vir  
Q40955 rubella vir  
Q991e7 rubella vir  
Q991e5 rubella vir  
Q9vhu8 drosophila  
Q9atk5 chlamydomon  
Q95996 homo sapien  
Q97298 plasmodium  
Q01677 bombyx mori  
Q60588 homo sapien  
Q14812 homo sapien  
Q9H450 homo sapien  
Q53393 mycobacteri  
Q9R0M0 mus musculu  
Q65553 bovine herp  
Q9P906 xylella fas  
Q45675 bacillus su  
Q97592 canis famil  
Q9PQ08 ureaplasma  
Q9VH01 drosophila  
Q9N9M2 leishmania

## ALIGNMENTS

RESULT 1  
Q9H7L3 PRELIMINARY; PRT; 477 AA.  
ID Q9H7L3  
AC Q9H7L3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE FLJ00059 PROTEIN (FRAGMENT).  
GN FLJ00059.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]

RN SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RA Ohara O., Nagase T., Kikuno R., Okumura K.;  
RT "The nucleotide sequence of a long cDNA clone isolated from human  
RT spleen.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK024466; BAB15756.1;  
FT NON\_TER 1  
SQ SEQUENCE 477 AA; 50773 MW; EFE3BE34BD0E1B05 CRC64;

Query Match 5.6%; Score 9; DB 4; Length 477;  
Best Local Similarity 100.0%; Pred. No. 0.88;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 DPARYPRYL 67  
ID 442 DPARYPRYL 450  
AC 442 DPARYPRYL 450

RESULT 2  
Q9H7Q8 PRELIMINARY; PRT; 540 AA.  
ID Q9H7Q8  
AC Q9H7Q8;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE FLJ00009 PROTEIN (FRAGMENT).  
GN FLJ00009.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RA Ohara O., Nagase T., Kikuno R., Okumura K.;  
RT "The nucleotide sequence of a long cDNA clone isolated from human  
RT spleen.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK024420; BAB15710.1;  
FT NON\_TER 1  
SQ SEQUENCE 540 AA; 57690 MW; 7776F207F28C426 CRC64;

Query Match 5.6%; Score 9; DB 4; Length 540;  
Best Local Similarity 100.0%; Pred. No. 0.98;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 DPARYPRYL 67  
ID 505 DPARYPRYL 513  
AC 505 DPARYPRYL 513

RESULT 3

Q9H7I6 PRELIMINARY; PRT; 672 AA.  
ID Q9H7I6  
AC Q9H7I6;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE FLJ00100 PROTEIN (FRAGMENT).  
GN FLJ00100.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RA Ohara O., Nagase T., Kikuno R., Okumura K.;  
RT "The nucleotide sequence of a long cDNA clone isolated from human  
RT spleen.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK024494; BAB15784.1;  
DR InterPro; IPR002293; AA\_rel\_permease\_1.  
FT NON\_TER 1  
SQ SEQUENCE 672 AA; 71622 MW; 1B27F8EDE38CDFDE CRC64;

Query Match 5.6%; Score 9; DB 4; Length 672;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 DPARYPRYL 67  
ID 637 DPARYPRYL 645  
AC 637 DPARYPRYL 645

RESULT 4

Q9H7Q7 PRELIMINARY; PRT; 772 AA.  
ID Q9H7Q7  
AC Q9H7Q7;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE FLJ00010 PROTEIN (FRAGMENT).  
GN FLJ00010.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

```

RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK024421; BAB15711.1; -.
DR InterPro; IPR002293; AA_rel_permease_1.
FT NON_TER 1
SQ SEQUENCE 772 AA; 81918 MW; B2CD21A1779FE5ED CRC64;

Query Match 5.6%; Score 9; DB 4; Length 772;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 DPARYPRYL 67
Db 737 DPARYPRYL 745
|||||

RESULT 5
ID Q99NC8 PRELIMINARY; PRT; 913 AA.
AC Q99NC8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE CATION-CHLORIDE COTRANSPORTER 6.
GN CCC6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Inai M., Ishibashi K.;
RT "Molecular cloning of a new member of cation-chloride cotransporter.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023645; BAB40440.1; -.
SQ SEQUENCE 913 AA; 95861 MW; C65B3394F7822523 CRC64;

Query Match 5.6%; Score 9; DB 11; Length 913;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 DPARYPRYL 67
Db 878 DPARYPRYL 886
|||||

RESULT 6
ID Q9NR5 PRELIMINARY; PRT; 914 AA.
AC Q9NR5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CATION-CHLORIDE COTRANSPORTER-INTERACTING PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isering P., Caron L.;
RT "Human Cation-Cl Cotransporter Interacting Protein (hCIP1).";
RL J. Biol. Chem. 0:0-0(2000).
DR EMBL; AF284422; AAF88060.1; -.
DR InterPro; IPR002293; AA_rel_permease_1.
SQ SEQUENCE 914 AA; 96170 MW; 78526E5D4B545472 CRC64;

```

```

Query Match 5.6%; Score 9; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 DPARYPRYL 67
Db 879 DPARYPRYL 887
|||||

RESULT 7
ID Q9BY10 PRELIMINARY; PRT; 914 AA.
AC Q9BY10;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE CATION-CHLORIDE COTRANSPORTER 6.
GN HCCC6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishibashi K.;
RT "Molecular cloning of human cation chloride cotransporter 6.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB033284; BAB40456.1; -.
SQ SEQUENCE 914 AA; 96078 MW; 6327F4B530BC77EF CRC64;

Query Match 5.6%; Score 9; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 DPARYPRYL 67
Db 879 DPARYPRYL 887
|||||

RESULT 8
ID Q9BXP2 PRELIMINARY; PRT; 914 AA.
AC Q9BXP2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE CATION-CHLORIDE COTRANSPORTER-INTERACTING PROTEIN 1.
GN CIP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21138439; PubMed=11239002;
RA Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,
RA Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
RA Miller W., Koop B.F.;
RT "Comparative analysis of the gene-dense ACHE/TFR2 region on human
RT chromosome 7q22 with the orthologous region on mouse chromosome 5.";
RL Nucleic Acids Res. 29:1352-1365(2001).
DR EMBL; AF312032; AAK21008.1; -.
SQ SEQUENCE 914 AA; 96109 MW; 8096855B78078CA4 CRC64;

Query Match 5.6%; Score 9; DB 4; Length 914;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 59 DPARYPRYL 67
|||||

```

Db 879 DPARYPYL 887

RESULT 9  
Q99MR3 PRELIMINARY; PRT; 914 AA.  
AC Q99MR3; (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE CIP1.  
GN Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RA Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,  
RA Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,  
RA Miller W., Koop B.F.;  
RT "Comparative analysis of the gene-dense ACHE/TFR2 region on human  
RT chromosome 7q22 with the orthologous region on mouse chromosome 5."  
RL Nucleic Acids Res. 29:1352-1365(2001).  
DR EMBL: AF312033; AAK28822.1; -  
SQ SEQUENCE 914 AA; 96313 MW; AF766F0A788CBEA3 CRC64;

Query Match 5.6%; Score 9; DB 11; Length 914;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 DPARYPYL 67

|||||||

Db 879 DPARYPYL 887

RESULT 10  
Q9HX52 PRELIMINARY; PRT; 261 AA.  
AC Q9HX52;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE HYPOTHETICAL PROTEIN PA3964.  
GN PA3964.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.W.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen."  
RL Nature 406:959-964(2000).  
DR EMBL: AE004814; AG07351.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 261 AA; 26810 MW; 85B9233DC5E6C36F CRC64;

Query Match 5.0%; Score 8; DB 2; Length 261;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 AGGRPADR 38  
Db 80 AGGRPADR 87  
|||||||

RESULT 11  
Q9UFZ6 PRELIMINARY; PRT; 804 AA.  
AC Q9UFZ6;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE HYPOTHETICAL 92.3 KDA PROTEIN (FRAGMENT).  
GN DKF2P434A014.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=TESTIS;  
RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL110224; CAB53682.1; -  
DR InterPro: IPR000130; Zn.MTpeptidse.  
DR InterPro: IPR001548; Peptidase\_M2.  
DR Pfam: PF01401; Peptidase\_M2; 1.  
DR PRINTS: PR00791; PEPDIPITASEA.  
DR ProDom: PD004184; Peptidase\_M2; 1.  
DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 804 AA; 92340 MW; 91FF391074CB5DA9 CRC64;

Query Match 5.0%; Score 8; DB 4; Length 804;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LFGEEDVR 88  
Db 663 LFGEEDVR 670  
|||||||

RESULT 12  
Q9NRA7 PRELIMINARY; PRT; 805 AA.  
AC Q9NRA7;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE ANGIOTENSIN CONVERTING ENZYME-LIKE PROTEIN (ACE-RELATED  
DE CARBOXYPEPTIDASE ACE2).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LYMPHOMA;  
RA Tipnis S.R., Hooper N.M., Hyde R.J., Christie G., Karran E.,  
RA Turner A.J.;  
RT "A Human Homolog of Angiotensin Converting Enzyme - Cloning and  
RT Functional Expression As A Captopril-Insensitive Carboxypeptidase."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Donoghue M., Hsieh F., Baronas E., Godbout K., Gosselin M.,  
RA Stagliano N., Donovan M., Woolf B., Robison K., Jeyaseelan R.,  
RA Breitbart R.E., Acton S.;  
RT "A novel ACE-related carboxypeptidase (ACE2) converts angiotensin I to  
RT angiotensin1-9.";

RL Circ. Res. 0:0-0(2000).  
 DR EMBL; AF2411254; AAF78220.1; -;  
 DR EMBL; AF2911820; AAF99721.1; -;  
 DR InterPro; IPR001548; Peptidase\_M2.  
 DR InterPro; IPR001130; Zn\_MTPeptide.  
 DR Pfam; PF01401; Peptidase\_M2; 1.  
 DR PRINTS; PR00791; PEPDPTASEA.  
 DR PRODOM; PD004184; Peptidase\_M2; 1.  
 DR POSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 KW Carboxypeptidase.  
 SQ SEQUENCE 805 AA; 92462 MW; 8EE6EBOA931550E8 CRC64;

Query Match 5.0%; Score 8; DB 4; Length 805;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LFGEEDVR 88  
 Db 664 LFGEEDVR 671

RESULT 13  
 Q9BYF1  
 ID Q9BYF1 PRELIMINARY; PRT; 805 AA.  
 AC Q9BYF1;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ACE2.  
 GN ACE2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Komatsu T., Suzuki Y., Sugano S.;  
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Suzuki Y., Watanabe M., Sugano S.;  
 RT "Cloning, expression analysis and chromosomal localization of a novel ACE like enzyme."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB046569; BAB40370.1; -;  
 SQ SEQUENCE 805 AA; 92491 MW; D2AAB4C27088EB72 CRC64;

Query Match 5.0%; Score 8; DB 4; Length 805;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LFGEEDVR 88  
 Db 664 LFGEEDVR 671

RESULT 14  
 Q9GIW8  
 ID Q9GIW8 PRELIMINARY; PRT; 47 AA.  
 AC Q9GIW8;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE MHC CLASS II ANTIGEN (FRAGMENT).  
 GN DOB.  
 OS Monachus schauinslandi (Hawaiian monk seal).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Monachus.  
 OX NCBI\_TaxID=29088;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=TSX3;  
 RA Armstrong P.S.;  
 RT "Low variation at the DQ loci of the Hawaiian monk seal.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- MISCELLANEOUS: FOUND N-TERMINAL TO IMMUNOGLOBULIN AND MAJOR  
 CC HISTOCOMPATIBILITY COMPLEX DOMAIN (BY SIMILARITY).  
 CC -!- SIMILARITY: TO CLASS II HISTOCOMPATIBILITY ANTIGEN, BETA CHAIN,  
 CC BETA-1 DOMAIN INTERPRO FAMILY.  
 DR EMBL; AY007204; AAG15838.1; -;  
 DR InterPro; IPR000353; MHC\_II\_beta.  
 DR Pfam; PF00969; MHC\_II\_beta; 1.  
 DR PRODOM; PD000328; MHC\_II\_beta; 1.  
 KW Glycoprotein; MHC; MHC II; Transmembrane.  
 FT NON\_TER 1 1  
 FT NON\_TER 47 47  
 SQ SEQUENCE 47 AA; 5435 MW; 398A3050B6375CA9 CRC64;

Query Match 4.4%; Score 7; DB 7; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 GEEDVRF 89  
 Db 21 GEEDVRF 27

RESULT 15  
 Q9PG71  
 ID Q9PG71 PRELIMINARY; PRT; 56 AA.  
 AC Q9PG71;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE HYPOTHETICAL PROTEIN XF0431.  
 GN XF0431.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 OX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA de Souza A.P., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
 RL Nature 406:151-159(2000).  
 DR EMBL; AB003894; AAF83241.1; -;

```

DR InterPro; IPR001128; Cyt_P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 56 AA; 6242 MW; 9289870E868DE3B CRC64;

Query Match
Best Local Similarity 4.4%; Score 7; DB 2; Length 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 DVRPRA 92
Db 31 DVRPRA 37

RESULT 16
Q9KW34 ID Q9KW34 PRELIMINARY; PRT; 92 AA.
AC Q9KW34;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE LYTIC ENZYME (FRAGMENT).
GN LYS.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SLM6;
RA Nakayama K., Takashima K., Ishihara H., Shinomiya S., Kageyama M.,
RA Kanaya S., Ohnishi M., Murata T., Mori H., Hayashi T.;
RT "The R-type pyocin is related to P2 phage, and the F-type pyocin is
RT related to lambda phage."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB045308; BAA97447.1; -.
FT NON_TER 1
FT SEQUENCE 92 AA; 10345 MW; B81FA9BC6F1D483B CRC64;

Query Match
Best Local Similarity 4.4%; Score 7; DB 2; Length 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RARAVLS 10
Db 86 RARAVLS 92

RESULT 17
Q9VX98 ID Q9VX98 PRELIMINARY; PRT; 131 AA.
AC Q9VX98;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG9099 PROTEIN.
GN CG9099.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.I., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Nusskern D.R., Pacieb J.M.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000)
DR EMBL; AE003503; AAR48679.2; -.
DR FlyBase; FBgn0030802; CG9099.
DR InterPro; IPR001950; SU11.
DR PROSITE; PS0296; SU11_2; 1.
SQ SEQUENCE 131 AA; 15024 MW; F0A3D59359FFD377 CRC64;

Query Match
Best Local Similarity 4.4%; Score 7; DB 5; Length 131;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LQLGPRE 22
Db 15 LQLGPRE 21

RESULT 18
Q9YCT7 ID Q9YCT7 PRELIMINARY; PRT; 150 AA.
AC Q9YCT7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHEICAL 15.5 KDA PROTEIN APE1175.
GN APE1175.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaei A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic

```

RT crenarchaeon, Aeropyrum pernix K1.;  
 RL DNA Res. 6:83-101(1999).  
 DR EMBL; AP000061; BAA80160.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 150 AA; 15467 MW; EAABFEA9CB6FCD83 CRC64;

Query Match 4.4%; Score 7; DB 1; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 AGRRPAD 37  
 | | | | | | |  
 Db 27 AGRRPAD 33

RESULT 19  
 O29550 PRELIMINARY; PRT; 154 AA.  
 ID AC O29550;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE CONSERVED HYPOTHETICAL PROTEIN.  
 GN AF0708.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 OC Archaeoglobus.  
 OX NCBI\_TaxID=2234;  
 RN [1]

SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49538;  
 RX MEDLINE=98049343; PubMed=9389475;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
 RA Cotton M.D., Spriggs T., Artlach P., Kaine B.P., Sykes S.M.,  
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon Archaeoglobus fulgidus."  
 RL Nature 390:364-370(1997).  
 DR EMBL; AE001056; AAB90533.1; -.  
 DR TIGR; AF0708; -.  
 DR InterPro: IPR002478; PUA.  
 DR Pfam; PF01472; PUA; 1.  
 DR SMART; SM00359; PUA; 1.  
 DR KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 154 AA; 17129 MW; 64B3B8AF27581AEF CRC64;

Query Match 4.4%; Score 7; DB 1; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 RAVLSAF 12  
 | | | | | | |  
 Db 129 RAVLSAF 135

RESULT 20  
 Q9RFQ8 PRELIMINARY; PRT; 154 AA.  
 ID AC Q9RFQ8;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE DNA LIGASE (EC 6.5.1.2) (FRAGMENT).

OS Rhodothermus marinus.  
 OC Bacteria; CFB group; Rhodothermus group; Rhodothermus.  
 OX NCBI\_TaxID=29549;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 4252;  
 RX MEDLINE=20054412; PubMed=10585410;  
 RA Martins L.O., Empadinhas N., Marugg J.D., Miguel C., Ferreira C.,  
 RA da Costa M.S., Santos H.;  
 RT "Biosynthesis of Mannosylglycerate in the Thermophilic Bacterium  
 RT Rhodothermus marinus. Biochemical and Genetic Characterization of a  
 RT Mannosylglycerate Synthase.";  
 RL J. Biol. Chem. 274:35407-35414(1999).  
 DR EMBL; AF173987; AAF16907.1; -.  
 DR HSSP; O87703; 1B04.  
 DR InterPro: IPR001679; DNA\_ligase\_N.  
 DR Pfam; PF01653; DNA\_ligase\_N; 1.  
 DR ProDom; PD003944; DNA\_ligase\_N; 1.  
 DR SMART; SM00532; LIGANC; 1.  
 KW Ligase.  
 FT NON\_TER 154 154  
 SQ SEQUENCE 154 AA; 17653 MW; 9FCCC085F2A1830B CRC64;

Query Match 4.4%; Score 7; DB 2; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 FGEEDYR 88  
 | | | | | | |  
 Db 120 FGEEDYR 126

RESULT 21  
 O26071 PRELIMINARY; PRT; 170 AA.  
 ID AC O26071;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE HYPOTHETICAL 19.6 KDA PROTEIN.  
 GN HP1546.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26695 / ATCC 700392;  
 RX MEDLINE=97394467; PubMed=9252185;  
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the gastric pathogen Helicobacter  
 RT pylori."  
 RL Nature 388:539-547(1997).  
 DR EMBL; AE000652; AAD08590.1; -.  
 DR TIGR; HP1546; -.  
 DR KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 170 AA; 19596 MW; 445831505EBA9308 CRC64;

Query Match 4.4%; Score 7; DB 2; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 132 EKDADSI 138



DE LYTIC ENZYME (HYPOTHETICAL PROTEIN PA0629).  
GN LYS OR PA0629.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RA Nakayama K., Takashima K., Ishihara H., Shinomiya T., Kageyama M.,  
RA Kanaya S., Ohnishi M., Murata T., Terawaki Y., Mori H., Hayashi T.;  
RT "Genetic relationship between bacteriocins and bacteriophages";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen";  
RL Nature 406:959-964(2000).  
DR EMBL; AB030825; BAA83168.1; -.  
DR EMBL; AE004498; AAG04018.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 209 AA; 23003 MW; BDD673C5D36A2B0E CRC64;

Query Match 4.4%; Score 7; DB 2; Length 209;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ARAVLS 10  
| | | | |  
Db 203 RARAVLS 209

RESULT 27  
P96378  
ID P96378 PRELIMINARY; PRT; 243 AA.  
AC P96378;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE HYPOTHETICAL 25.8 KDA PROTEIN.  
GN LP00 OR RV1022 OR MTC10G2.27C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaija F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,  
RA Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence";  
RL Nature 393:537-544(1998).  
DR EMBL; Z92539; CAB06857.1; -.  
DR TuberculList; Rv1022; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 243 AA; 25932 MW; 84246C42CAF7EF06 CRC64;

Query Match 4.4%; Score 7; DB 2; Length 243;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 AGRRPAD 37  
| | | | |  
Db 51 AGRRPAD 57

RESULT 28  
Q9K3H8  
ID Q9K3H8 PRELIMINARY; PRT; 253 AA.  
AC Q9K3H8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE PUTATIVE GNTFR-FAMILY TRANSCRIPTIONAL REGULATOR.  
GN 2SCG18.09C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Brown S.P., Harris D.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kleser H.M., Denapaita D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome";  
RL Mol. Microbiol. 21:77-96(1996).  
CC -1- SIMILARITY: BELONGS TO THE GNTFR FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS.  
DR EMBL; AL390188; CAB99142.1; -.  
DR InterPro; IPR000524; HTH\_GntR.  
DR Pfam; PF00392; gntR; 1.  
DR PRINTS; PR00035; HTHGNTFR.  
DR SMART; SM00345; HTH\_GNTFR; 1.  
KW DNA-binding; Transcription regulation.  
SQ SEQUENCE 253 AA; 27143 MW; 89BBBFC5B497AC3C CRC64;

Query Match 4.4%; Score 7; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ARARAVL 9  
| | | | |  
Db 164 ARARAVL 170

RESULT 29  
O50480  
ID O50480 PRELIMINARY; PRT; 288 AA.  
AC O50480;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE PUTATIVE TRANSCRIPTIONAL REGULATOR.  
GN SC7C7.17 OR SC4H8.01.  
OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 RX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Harris D., Taylor K.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Redenbach M., Kleser H.M., Denapaita D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]  
 RP SEQUENCE OF 260-288 FROM N.A.  
 RC STRAIN=A3(2);  
 RA Harris D.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 260-288 FROM N.A.  
 RC STRAIN=A3(2);  
 RA Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE ARAC/XLYS FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 DR EMBL; AL031031; CAA19863.1; -;  
 DR EMBL; AL020958; CAA15868.1; -;  
 DR InterPro; IPR000005; HTHARAC.  
 DR Pfam; PF00165; HTH\_Arac; 1.  
 DR SMART; SM00342; HTH\_ARAC; 1.  
 DR PROSITE; PS00041; HTH\_ARAC\_FAMILY\_1; FALSE\_NEG.  
 DR PROSITE; PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
 KW DNA-binding; Hypothetical protein; Transcription regulation.  
 SQ SEQUENCE 288 AA; 31363 MW; 8BAF85286D73CB0B CRC64;

Query Match 4.4%; Score 7; DB 2; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ARARAVL 9  
 DB 174 ARARAVL 180

RESULT 30  
 ID P96356 PRELIMINARY; PRT; 293 AA.  
 AC P96356;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE HYPOTHETICAL 32.0 KDA PROTEIN.  
 GN RV1045 OR MTCY10G2.04C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,  
 RA Taylor K., Whitehead S., Barrell B.G.;  
 RT "Completing the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 DR EMBL; Z92539; CAB08867.1; -;  
 DR Tuberculist; RV1045; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 293 AA; 31990 MW; 3CEC83C346F5CB17 CRC64;

Query Match 4.4%; Score 7; DB 2; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 PPTNLRs 48  
 DB 8 PPTNLRs 14

RESULT 31  
 ID 026103 PRELIMINARY; PRT; 307 AA.  
 AC 026103;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE PYRIDOXAL PHOSPHATE BIOSYNTHETIC PROTEIN A (PDXA).  
 GN HP1583  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26695 / ATCC 700392;  
 RX MEDLINE=97394467; PubMed=9252185;  
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the gastric pathogen Helicobacter  
 RT pylori.";  
 RL Nature 388:539-547(1997).  
 DR EMBL; AE000655; AAD08621.1; -;  
 DR TIGR; HP1583; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 307 AA; 33578 MW; B12280A0BB00D70E CRC64;

Query Match 4.4%; Score 7; DB 2; Length 307;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GLFGEED 86  
 DB 200 GLFGEED 206

RESULT 32  
 ID 092J28 PRELIMINARY; PRT; 307 AA.  
 AC 092J28;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)

```

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PRIDOXAL PHOSPHATE BIOSYNTHETIC PROTEIN A.
GN PDXA OR JHP1490.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL; AE001570; AAD07064.1; -.
KW Complete proteome.
SQ SEQUENCE 307 AA; 33706 MW; 77C5FD2DD9A71A81 CRC64;

Query Match 4.4%; Score 7; DB 2; Length 307;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GLFGED 86
| | | | |
DB 200 GLFGED 206

RESULT 33
P72344 PRELIMINARY; PRT; 331 AA.
ID P72344
AC P72344;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ORF993.
OS Rhodobacter sphaeroides (Rhodospirillum rubrum sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=WS8;
RA Pasternak C., Assemet K., Clement-Metral J.D., Klug G.;
RT "Thioredoxin is essential for Rhodobacter sphaeroides growth by
RT aerobic and anaerobic respiration.";
RL Microbiology 143:83-91(1997).
DR EMBL; U69571; AAB51777.1; -.
SQ SEQUENCE 331 AA; 35768 MW; FFD4389792ABB650 CRC64;

Query Match 4.4%; Score 7; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ARAVLSA 11
| | | | |
DB 224 ARAVLSA 230

RESULT 34
Q92538 PRELIMINARY; PRT; 332 AA.
ID Q92538
AC Q92538;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

```

```

DE PUTATIVE REGULATORY PROTEIN.
GN SC9B2.21C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmid and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL035212; CAA22801.1; -.
DR InterPro; IPR000523; Mg_chelatase_chII.
DR Pfam; PF01078; Mg_chelatase; 1.
SQ SEQUENCE 332 AA; 36001 MW; 621E19AB291B1DAE CRC64;

Query Match 4.4%; Score 7; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ARARAVL 9
| | | | |
DB 266 ARARAVL 272

RESULT 35
Q9YUR7 PRELIMINARY; PRT; 367 AA.
ID Q9YUR7
AC Q9YUR7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE IVA2 PROTEIN.
GN IVA2.
OS turkey adenovirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Siadenovirus.
OX NCBI_TaxID=41678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99009384; PubMed=9791022;
RA Pitcovski J., Muallem M., Rei-Koren Z., Krisspel S., Gallili G.,
RA Michael A., Goldberg D.;
RT "The complete DNA sequence and genome organization of the avian
RT adenovirus, hemorrhagic enteritis virus.";
RL Virology 249:307-315(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Pitcovski J., Muallem M., Rei-Koren Z., Krisspel S., Gallili G.,
RA Michael A., Goldberg D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF074946; AAC64524.1; -.
DR InterPro; IPR001687; ATP-GTP_A.
DR Pfam; PF02456; Adeno_IVA2; 1.
SQ SEQUENCE 367 AA; 41803 MW; C7C9F88C86D3937C CRC64;

Query Match 4.4%; Score 7; DB 12; Length 367;

```

```

Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 DSINSSI 142
    |||||
Db 20 DSINSSI 26

RESULT 36
O16765 PRELIMINARY; PRT; 399 AA.
AC O16765;
AT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DR 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE R07C3.11 PROTEIN.
GN R07C3.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RN Nature 368:32-38(1994).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Lamar B., Kramer J.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL EMBL; AF016686; AAB66240.1; -.
SQ SEQUENCE 399 AA; 46626 MW; A8D6865596839E64 CRC64;

Query Match 4.4%; Score 7; DB 5; Length 399;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 ISYDPAR 62
    |||||
Db 380 ISYDPAR 386

RESULT 37
Q9NFY2 PRELIMINARY; PRT; 435 AA.
AC Q9NFY2;
AT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DR 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SERINE PROTEASE.
GN SER1.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
```

```

Anopheles.
OC NCBI_TaxID=71165;
RN [1]
RP SEQUENCE FROM N.A.
RA Volz J.B.C., Kafatos F.C., Muller H.M.;
RT "PCR-Cloning of a serine protease AgSer1 in Anopheles gambiae.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC -I- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (SL).
DR EMBL; AJ250916; CAB91840.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Trypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp-SPC; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 435 AA; 48448 MW; 2DEAC78923B04B61 CRC64;

Query Match 4.4%; Score 7; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 AGGRPAD 37
    |||||
Db 204 AGGRPAD 210

RESULT 38
Q51006 PRELIMINARY; PRT; 467 AA.
ID Q51006
AC Q51006;
AT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DR 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MTRF GENE (FRAGMENT).
GN MTRF.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FAL36;
RX MEDLINE=97388568; PubMed=9245802;
RA Delahay R.M., Robertson B.D., Bathazar J.T., Shafer W.M., Ison C.A.;
RT "Involvement of the gonococcal MtrF protein in the resistance of
RL Neisseria gonorrhoeae to toxic hydrophobic agents.";
RL Microbiology 143:2127-2133(1997).
DR EMBL; X95635; CAA64891.1; -.
DR InterPro; IPR003423; OEP.
DR Pfam; PF02321; OEP; 2.
FT NON_TER 467 467
SQ SEQUENCE 467 AA; 50463 MW; E92B7DE5FEA2EF14 CRC64;

Query Match 4.4%; Score 7; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 REQARNA 27
    |||||
Db ~ 235 REQARNA 241

RESULT 39
Q9JY68 PRELIMINARY; PRT; 467 AA.
ID Q9JY68
AC Q9JY68;
AT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DR 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```

DE MULTIDRUG EFFLUX PUMP CHANNEL PROTEIN.  
GN NMB1714.  
OS Neisseria meningitidis (serogroup B).  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MC58 / SEROGROUP B;  
RX MEDLINE=20175755; PubMed=10710307;  
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,  
Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
Mason T., Clecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,  
Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,  
Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,  
Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;  
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain  
MC58.";  
RL Science 287:1809-1815(2000).  
DR EMBL: AE002521; AAF42061.1; -.  
DR TIGR: NMB1714; -.  
DR InterPro: IPR003423; OEP.  
DR Pfam: PF02321; OEP; 2.  
KW Complete proteome.  
SQ SEQUENCE 467 AA; 50476 MW; C76032302F34F939 CRC64;

Query Match 4.4%; Score 7; DB 2; Length 467;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 REQARNA 27  
Db 235 REQARNA 241  
|||||

RESULT 40  
Q9JT52 ID Q9JT52 PRELIMINARY; PRT; 467 AA.  
AC Q9JT52;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE PUTATIVE OUTER MEMBRANE LIPOPROTEIN.  
GN MRE OR NMA1968  
OS Neisseria meningitidis (serogroup A).  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=65699;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;  
RX MEDLINE=2022556; PubMed=10761919;  
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,  
Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holroyd S.,  
Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,  
Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
Whitehead S., Spratt B.G., Barrall B.G.;  
RT "Complete DNA sequence of a serogroup A strain of Neisseria  
meningitidis Z2491.";  
RL Nature 404:502-506(2000).  
DR EMBL: AL162757; CAB85188.1; -.  
DR InterPro: IPR003423; OEP.  
DR Pfam: PF02321; OEP; 2.  
KW Lipoprotein; Complete proteome.  
SQ SEQUENCE 467 AA; 50614 MW; 9E40154E2BF31649 CRC64;

Query Match 4.4%; Score 7; DB 2; Length 467;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 REQARNA 27  
Db 235 REQARNA 241  
|||||

RESULT 41  
Q9P8G5 ID Q9P8G5 PRELIMINARY; PRT; 470 AA.  
AC Q9P8G5;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE PUTATIVE SNARE PROTEIN.  
GN SYN.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gupta G.D., Heath B.I.;  
RT "Neurospora crassa putative SNARE, similar to yeast SSOlp.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF262964; AAF72704.1; -.  
DR InterPro: IPR000017; Syntaxin.  
DR InterPro: IPR000727; T-SNARE.  
DR Pfam: PF00804; Syntaxin; 1.  
DR SMART: SM00503; Synn; 1.  
DR SMART: SM00397; T-SNARE; 1.  
SQ SEQUENCE 470 AA; 52175 MW; 2FCB4CFECC60F934 CRC64;

Query Match 4.4%; Score 7; DB 3; Length 470;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ARAVLSA 11  
Db 214 ARAVLSA 220  
|||||

RESULT 42  
Q9KHC4 ID Q9KHC4 PRELIMINARY; PRT; 486 AA.  
AC Q9KHC4;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE SOCE.  
GN SOCE.  
OS Myxococcus xanthus.  
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.  
OX NCBI\_TaxID=34;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20156361; PubMed=10691740;  
RA Crawford E.W. Jr., Shimkets L.J.;  
RT "The stringent response in Myxococcus xanthus is regulated by SocE and  
the csgA C-signaling protein.";  
RL Genes Dev. 14:483-492(2000).  
DR EMBL: AF263243; AAF91388.1; -.  
SQ SEQUENCE 486 AA; 53554 MW; 3D5BE4F0F0746038 CRC64;

Query Match 4.4%; Score 7; DB 2; Length 486;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ACAGRS 112  
Db 193 ACAGRS 199  
|||||

```
RESULT 43
Q13512 ID Q13512 PRELIMINARY; PRT; 551 AA.
AC Q13512
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE PROTEIN B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FIBROSCOMA; PubMed=8723724;
RX MEDLINE=96303695;
RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
  Malley T., Gibbs R.A.;
RT "A gene-rich cluster between the CD4 and triosephosphate isomerase
  genes at human chromosome 12p13."
RL Genome Res. 6:314-326(1996).
DR ENBL; U47926; AAC50464.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 551 AA; 62264 MW; F16E4048C0664F58 CRC64;

Query Match 4.4%; Score 7; DB 4; Length 551;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 QGAKLLL 151
DB 345 QGAKLLL 351

RESULT 44
Q15740 ID Q15740 PRELIMINARY; PRT; 551 AA.
AC Q15740
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96303695; PubMed=8723724;
RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,
  Malley T., Gibbs R.A.;
RT "A gene-rich cluster between the CD4 and triosephosphate isomerase
  genes at human chromosome 12p13."
RL Genome Res. 6:314-326(1996).
DR ENBL; U47924; AAB51312.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 551 AA; 62294 MW; 1CAA483E15659886 CRC64;
```

```
Query Match 4.4%; Score 7; DB 4; Length 551;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 QGAKLLL 151
DB 345 QGAKLLL 351
```

```
RESULT 45
Q9UDU4 ID Q9UDU4 PRELIMINARY; PRT; 699 AA.
AC Q9UDU4
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE WUGSC:H.DJ1140G11.1 PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Du H., Wollam C.;
RT "The sequence of Homo sapiens PAC clone RP5-1140G11."
RN [2]
RP Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RA Waterston R.H.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR ENBL; AC004974; AAC83179.1; -.
DR InterPro; IPR000331; Rap_GAP.
FT NON_TER 699
SQ SEQUENCE 699 AA; 77909 MW; 1DE9A632C23B4507 CRC64;

Query Match 4.4%; Score 7; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 SSIDKQG 146
DB 210 SSIDKQG 216

RESULT 46
Q9ASG3 ID Q9ASG3 PRELIMINARY; PRT; 812 AA.
AC Q9ASG3
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE P0686E09.18 PROTEIN.
GN P0686E09.18
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
  clone:P0686E09."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL; AP002897; BAB39125.1; -.
SQ SEQUENCE 812 AA; 90595 MW; B5B4F940E01BA1E2 CRC64;
```

```
Query Match 4.4%; Score 7; DB 10; Length 812;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 GEEDVRF 89
DB 596 GEEDVRF 602
```

RESULT 47  
O43701 PRELIMINARY; PRT; 830 AA.  
AC O43701;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE ACETYL LDL RECEPTOR PRECURSOR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98058897; PubMed=9395444;  
RA Adachi H., Tsujimoto M., Arai H., Inoue K.;  
RT "Expression cloning of a novel scavenger receptor from human  
RT endothelial cells.";  
RL J. Biol. Chem. 272:31217-31220(1997).  
DR EMBL; D86864; BAA24070.1; -.  
DR HSSP; P00742; 1XKA.  
DR InterPro; IPR000561; EGF-like.  
DR SMART; SM00181; EGF; 5.  
DR PROSITE; PS00001; EGF\_like; 2.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_6.  
DR PROSITE; PS01186; EGF\_2; 6.  
KW EGF-like domain; Glycoprotein; Signal.  
FT SIGNAL 1 15 POTENTIAL.  
FT CHAIN 16 830 ACETYL LDL RECEPTOR.  
SQ SEQUENCE 830 AA; 87430 MW; F560D9E1AA64D779 CRC64;

Query Match 4.4%; Score 7; DB 4; Length 830;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 NASCPAG 32  
Db 382 NASCPAG 388

RESULT 48  
Q14162 PRELIMINARY; PRT; 830 AA.  
AC Q14162;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE KIAA0149 PROTEIN.  
GN KIAA0149.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96127530; PubMed=8590280;  
RA Nagase T., Seki N., Tanaka A., Ishikawa K., Nomura N.;  
RT "Prediction of the coding sequences of unidentified human genes. IV.  
RT The coding sequences of 40 new genes (K1AA0121-K1AA0160) deduced by  
RT analysis of cDNA clones from human cell line KG-1.";  
RL DNA Res. 2:167-174(1995).  
DR EMBL; D63483; BAA09770.1; -.  
DR HSSP; P35555; 1EMN.  
DR InterPro; IPR000561; EGF-like.  
DR SMART; SM00181; EGF; 5.  
DR PROSITE; PS00001; EGF\_like; 2.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_6.  
DR PROSITE; PS01186; EGF\_2; 6.  
KW EGF-like domain; Glycoprotein.  
SQ SEQUENCE 830 AA; 87460 MW; 81748A694B25D36D CRC64;

Query Match 4.4%; Score 7; DB 4; Length 830;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 26 NASCPAG 32  
Db 382 NASCPAG 388

RESULT 49  
Q69139 PRELIMINARY; PRT; 946 AA.  
AC Q69139;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE NUCLEAR ANTIGEN EBNA-3B.  
OS Human herpesvirus 4 (Epstein-Barr virus).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gammaherpesvirinae; Lymphocryptovirus.  
OX NCBI\_TaxID=10376;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90347798; PubMed=2166806;  
RA Sample J., Young L., Martin B., Chatman T., Kieff E.D., Rickinson A.;  
RT "Epstein-Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B,  
RT and EBNA-3C genes.";  
RL J. Virol. 64:4084-4092(1990).  
DR EMBL; M34440; AAA45894.1; -.  
SQ SEQUENCE 946 AA; 104103 MW; 7F428D409134E869 CRC64;

Query Match 4.4%; Score 7; DB 12; Length 946;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 PTVVLR 103  
Db 450 PTVVLR 456

RESULT 50  
O44066 PRELIMINARY; PRT; 966 AA.  
AC O44066;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE DNA POLYMERASE ALPHA (FRAGMENT).  
OS Paraurostyla viridis.  
OC Eukaryota; Alveolata; Ciliophora; hypotrichs; Stichotrichida;  
OC Amphiseliidae; Paraurostyla.  
OX NCBI\_TaxID=57502;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97451043; PubMed=9302325;  
RA Hoffman D.C., Prescott D.M.;  
RT "Phylogenetic relationships among hypotrichous ciliates determined  
RT with the macronuclear gene encoding the large, catalytic subunit of  
RT DNA polymerase alpha.";  
RL J. Mol. Evol. 45:301-310(1997).  
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N  
CC PYROPHOSPHATE + DNA(N).  
CC -!- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:  
CC ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR  
CC DIFFERENT REACTIONS OF DNA SYNTHESIS (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.  
DR EMBL; U89701; AAC38856.1; -.  
DR InterPro; IPR002064; DNA\_pol\_B.  
DR Pfam; PF00136; DNA\_pol\_B; 1.  
DR PRINTS; PR00106; DNAPOLB.  
DR SMART; SM00486; POLBc; 1.

KW DNA replication; DNA-binding; DNA-directed DNA polymerase.  
FT NON\_TER 1  
FT NON\_TER 966  
SQ SEQUENCE 966 AA; 111627 MW; E472EFFDAD8407A6 CRC64;

Query Match 4.48; Score 7; DB 5; Length 966;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 GLFGGED 86  
|||||||  
Db 767 GLFGGED 773

Search completed: April 14, 2002, 12:45:51  
Job time: 213 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2002, 12:39:08 ; Search time 12.73 Seconds  
(without alignments)  
282.838 Million cell updates/sec

Title: US-09-731-816-4

Perfect score: 160

Sequence: 1 NSARARAVLSAFHHTLQLGP.....SIDKQAKLLGPNAPAGP 160

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 6

Total number of hits satisfying chosen parameters: 120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2.6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2.6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2.6/ptodata/2/iaa/6A\_COMB.pep.\*

4: /cgn2.6/ptodata/2/iaa/6B\_COMB.pep.\*

5: /cgn2.6/ptodata/2/iaa/PTUS\_COMB.pep.\*

6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	5.0	805	4	US-08-989-299-2
2	7	4.4	1580	2	US-08-804-227C-11
3	7	4.4	1580	2	US-08-804-198-5
4	6	3.8	13	6	5242798-11
5	6	3.8	79	4	US-08-432-994A-6
6	6	3.8	83	6	5242798-1
7	6	3.8	150	3	US-09-034-810-4
8	6	3.8	150	3	US-08-685-239-4
9	6	3.8	150	4	US-08-432-994A-2
10	6	3.8	151	2	US-08-620-694A-8
11	6	3.8	151	3	US-09-034-810-6
12	6	3.8	151	3	US-09-022-255-8
13	6	3.8	151	3	US-09-022-696-8
14	6	3.8	151	3	US-08-685-239-6
15	6	3.8	151	3	US-09-022-253-8
16	6	3.8	151	3	US-09-022-260-8
17	6	3.8	151	4	US-09-022-259-8
18	6	3.8	151	4	US-09-022-257-8
19	6	3.8	151	4	US-08-432-994A-4
20	6	3.8	151	1	US-08-514-014-12
21	6	3.8	153	2	US-08-833-823-12
22	6	3.8	155	4	US-08-432-994A-8
23	6	3.8	158	2	US-08-620-694A-7
24	6	3.8	158	3	US-09-022-255-7
25	6	3.8	158	3	US-09-022-696-7
26	6	3.8	158	3	US-09-022-253-7
27	6	3.8	158	3	US-09-022-260-7
28	6	3.8	158	3	US-09-022-259-7
29	6	3.8	158	3	US-09-022-257-7
30	6	3.8	158	3	US-09-022-257-7
31	6	3.8	158	3	US-09-022-257-7
32	6	3.8	158	3	US-09-022-257-7
33	6	3.8	158	3	US-09-022-257-7
34	6	3.8	158	3	US-09-022-257-7
35	6	3.8	158	3	US-09-022-257-7
36	6	3.8	158	3	US-09-022-257-7
37	6	3.8	158	3	US-09-022-257-7
38	6	3.8	158	3	US-09-022-257-7
39	6	3.8	158	3	US-09-022-257-7
40	6	3.8	158	3	US-09-022-257-7
41	6	3.8	158	3	US-09-022-257-7
42	6	3.8	158	3	US-09-022-257-7
43	6	3.8	158	3	US-09-022-257-7
44	6	3.8	158	3	US-09-022-257-7
45	6	3.8	158	3	US-09-022-257-7
46	6	3.8	158	3	US-09-022-257-7
47	6	3.8	158	3	US-09-022-257-7
48	6	3.8	158	3	US-09-022-257-7
49	6	3.8	158	3	US-09-022-257-7
50	6	3.8	158	3	US-09-022-257-7
51	6	3.8	158	3	US-09-022-257-7
52	6	3.8	158	3	US-09-022-257-7
53	6	3.8	158	3	US-09-022-257-7
54	6	3.8	158	3	US-09-022-257-7
55	6	3.8	158	3	US-09-022-257-7
56	6	3.8	158	3	US-09-022-257-7
57	6	3.8	158	3	US-09-022-257-7
58	6	3.8	158	3	US-09-022-257-7
59	6	3.8	158	3	US-09-022-257-7
60	6	3.8	158	3	US-09-022-257-7
61	6	3.8	158	3	US-09-022-257-7
62	6	3.8	158	3	US-09-022-257-7
63	6	3.8	158	3	US-09-022-257-7
64	6	3.8	158	3	US-09-022-257-7
65	6	3.8	158	3	US-09-022-257-7
66	6	3.8	158	3	US-09-022-257-7
67	6	3.8	158	3	US-09-022-257-7
68	6	3.8	158	3	US-09-022-257-7
69	6	3.8	158	3	US-09-022-257-7
70	6	3.8	158	3	US-09-022-257-7
71	6	3.8	158	3	US-09-022-257-7
72	6	3.8	158	3	US-09-022-257-7
73	6	3.8	158	3	US-09-022-257-7
74	6	3.8	158	3	US-09-022-257-7
75	6	3.8	158	3	US-09-022-257-7
76	6	3.8	158	3	US-09-022-257-7
77	6	3.8	158	3	US-09-022-257-7
78	6	3.8	158	3	US-09-022-257-7
79	6	3.8	158	3	US-09-022-257-7
80	6	3.8	158	3	US-09-022-257-7
81	6	3.8	158	3	US-09-022-257-7
82	6	3.8	158	3	US-09-022-257-7
83	6	3.8	158	3	US-09-022-257-7
84	6	3.8	158	3	US-09-022-257-7
85	6	3.8	158	3	US-09-022-257-7
86	6	3.8	158	3	US-09-022-257-7
87	6	3.8	158	3	US-09-022-257-7
88	6	3.8	158	3	US-09-022-257-7
89	6	3.8	158	3	US-09-022-257-7
90	6	3.8	158	3	US-09-022-257-7
91	6	3.8	158	3	US-09-022-257-7
92	6	3.8	158	3	US-09-022-257-7
93	6	3.8	158	3	US-09-022-257-7
94	6	3.8	158	3	US-09-022-257-7
95	6	3.8	158	3	US-09-022-257-7
96	6	3.8	158	3	US-09-022-257-7
97	6	3.8	158	3	US-09-022-257-7
98	6	3.8	158	3	US-09-022-257-7
99	6	3.8	158	3	US-09-022-257-7
100	6	3.8	158	3	US-09-022-257-7

101 6 3.8 1078 3 US-08-480-474-11 Sequence 11, Appl  
102 6 3.8 1138 1 US-07-973-320-2 Sequence 2, Appl  
103 6 3.8 1138 1 US-07-973-320-4 Sequence 4, Appl  
104 6 3.8 1205 4 US-09-352-159-29 Sequence 29, Appl  
105 6 3.8 1205 4 US-09-352-168-29 Sequence 29, Appl  
106 6 3.8 1213 4 US-09-413-814-79 Sequence 79, Appl  
107 6 3.8 1452 2 US-08-652-971-4 Sequence 4, Appl  
108 6 3.8 1452 2 US-08-449-644-8 Sequence 8, Appl  
109 6 3.8 1452 2 US-08-087-244A-8 Sequence 8, Appl  
110 6 3.8 1452 2 US-08-991-258A-4 Sequence 4, Appl  
111 6 3.8 1452 2 US-08-769-393-4 Sequence 4, Appl  
112 6 3.8 1452 3 US-08-991-953A-4 Sequence 4, Appl  
113 6 3.8 1464 1 US-08-026-138E-1 Sequence 1, Appl  
114 6 3.8 1464 2 US-08-231-193A-11 Sequence 11, Appl  
115 6 3.8 1464 2 US-08-486-273A-11 Sequence 11, Appl  
116 6 3.8 1464 3 US-08-940-086A-11 Sequence 11, Appl  
117 6 3.8 1464 4 US-08-436-332B-10 Sequence 10, Appl  
118 6 3.8 2205 1 US-08-093-453B-2 Sequence 2, Appl  
119 6 3.8 3567 2 US-07-642-734C-4 Sequence 4, Appl  
120 6 3.8 3567 3 US-08-439-009A-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-08-989-299-2  
; Sequence 2, Application US/08989299  
; Patent No. 6194556  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan L.  
; APPLICANT: Robinson, Keith E.  
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG  
; AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/989,299  
; FILING DATE: 11-DEC-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arnold E., Beth  
; REGISTRATION NUMBER: 35,430  
; REFERENCE/DOCKET NUMBER: MIA-025.01  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 805 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-989-299-2

Query Match 5.0%; Score 8; DB 4; Length 805;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 81 LFGEEDVR 88  
|

Db 664 LFGEEDVR 671  
RESULT 2  
US-08-804-227C-11  
; Sequence 11, Application US/08804227C  
; Patent No. 5876991  
; GENERAL INFORMATION:  
; APPLICANT: DeHoff, Bradley S.  
; APPLICANT: Kuhstoss, Stuart A.  
; APPLICANT: Rostock, Paul R., Jr.  
; APPLICANT: Sutton, Kimberly L.  
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: THOMAS G. PLANT 1501  
; STREET: LILLY CORPORATE CENTER  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: USA  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM Compatible  
; SOFTWARE: ASCII(DOS) Text only  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/804,227C  
; FILING DATE: February 21, 1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plant, Thomas, G.  
; REGISTRATION NUMBER: 35,784  
; REFERENCE/DOCKET NUMBER: X-8231  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-2459  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1580 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-804-227C-11  
Query Match 4.4%; Score 7; DB 2; Length 1580;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 30 PAGGRPA 36  
|  
Db 1518 PAGGRPA 1524  
RESULT 3  
US-08-804-198-5  
; Sequence 5, Application US/08804198  
; Patent No. 5945320  
; GENERAL INFORMATION:  
; APPLICANT: Burgett, Stanley G.  
; APPLICANT: Kuhstoss, Stuart A.  
; APPLICANT: Rao, Nagaraja R.  
; APPLICANT: Richardson, Mark A.  
; APPLICANT: Rostock, Paul R., Jr.  
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: PAUL R. CANTRELL 1138  
; STREET: LILLY CORPORATE CENTER  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: USA  
; ZIP: 46285

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh 7.0  
SOFTWARE: Microsoft Word 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804,198  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CANTRELL, PAUL R.  
REGISTRATION NUMBER: 36,470  
REFERENCE/DOCKET NUMBER: P9113  
TELEPHONE: 317-276-3885  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1580 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-804-198-5

Query Match 4.4%; Score 7; DB 2; Length 1580;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PAGGRPA 36  
Db 1518 PAGGRPA 1524

RESULT 4  
5242798-11  
; Patent No. 5242798  
; APPLICANT: SUTCLIFFE, J. GERGOR  
; TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES CORRESPONDING  
; TO PORTIONS OF PROTEINOIDS TRANSLATED FROM BRAIN-SPECIFIC MRNAS,  
; RECEPTORS, METHODS AND DIAGNOSTICS USING THE SAME  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/476,961  
; FILING DATE: 07-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 58,620  
; FILING DATE: 03-JUN-1987  
; APPLICATION NUMBER: 516,136  
; FILING DATE: 21-JUL-1983  
; SEQ ID NO:11:  
; LENGTH: 13  
5242798-11

Query Match 3.8%; Score 6; DB 6; Length 13;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 RSVSPW 52  
Db 1 RSVSPW 6

RESULT 5  
US-08-432-994A-6  
; Sequence 6, Application US/08432994A  
; Patent No. 6274711  
; GENERAL INFORMATION:  
; APPLICANT: Golstein, Pierre  
; APPLICANT: Rouvier, Eric  
; APPLICANT: Fossiez, Francois  
; APPLICANT: Lebecque, Serge J.E.  
; APPLICANT: Djossou, Odile

APPLICANT: Bancheureau, Jacques  
TITLE OF INVENTION: PURIFIED MAMMALIAN CTLA-8 ANTIGENS AND  
RELATED REAGENTS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/432,994A  
FILING DATE: 02-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/250,846  
FILING DATE: 27-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/177,747  
FILING DATE: 05-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,203  
FILING DATE: 14-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Egwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0388K3  
TELEPHONE: 415-852-9196  
TELEFAX: 415-496-1200  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 79 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-432-994A-6

Query Match 3.8%; Score 6; DB 4; Length 79;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
Db 66 VGCTCV 71

RESULT 6  
5242798-1  
; Patent No. 5242798  
; APPLICANT: SUTCLIFFE, J. GERGOR  
; TITLE OF INVENTION: SYNTHETIC POLYPEPTIDES CORRESPONDING  
; TO PORTIONS OF PROTEINOIDS TRANSLATED FROM BRAIN-SPECIFIC MRNAS,  
; RECEPTORS, METHODS AND DIAGNOSTICS USING THE SAME  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/476,961  
; FILING DATE: 07-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 58,620  
; FILING DATE: 03-JUN-1987  
; APPLICATION NUMBER: 516,136  
; FILING DATE: 21-JUL-1983  
; SEQ ID NO:1:  
; LENGTH: 83  
5242798-1

Query Match 3.8%; Score 6; DB 6; Length 83;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 RSVSPW 52  
Db 57 RSVSPW 62

## RESULT 7

US-09-034-810-4  
; Sequence 4, Application US/09034810  
; Patent No. 6043344  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: Kelleher, Kerry  
; APPLICANT: Carlin, McKeough  
; APPLICANT: Goldman, Samuel  
; APPLICANT: Pittman, Debra  
; APPLICANT: Mi, Sha  
; APPLICANT: Neben, Steven  
; APPLICANT: Giannotti, JoAnn  
; APPLICANT: Golden/Fleet, Margaret  
; TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/034,810  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/685,239  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brown, Scott A.  
; REFERENCE/DOCKET NUMBER: G15262  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8224  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 150 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-034-810-4

Query Match 3.8%; Score 6; DB 3; Length 150;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
Db 137 VGCTCV 142

## RESULT 8

US-08-685-239-4

; Sequence 4, Application US/08685239  
; Patent No. 6074849  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: Kelleher, Kerry  
; APPLICANT: Carlin, McKeough  
; APPLICANT: Goldman, Samuel  
; APPLICANT: Pittman, Debra  
; APPLICANT: Mi, Sha  
; APPLICANT: Neben, Steven  
; APPLICANT: Giannotti, JoAnn  
; APPLICANT: Golden/Fleet, Margaret  
; TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,239  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brown, Scott A.  
; REFERENCE/DOCKET NUMBER: G15262  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8224  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 150 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-685-239-4

Query Match 3.8%; Score 6; DB 3; Length 150;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
Db 137 VGCTCV 142

## RESULT 9

US-08-432-994A-2  
; Sequence 2, Application US/08432994A  
; Patent No. 6274711  
; GENERAL INFORMATION:  
; APPLICANT: Golstein, Pierre  
; APPLICANT: Rouvier, Eric  
; APPLICANT: Fossiez, Francois  
; APPLICANT: Lebecque, Serge J.E.  
; APPLICANT: Djossou, Odile  
; APPLICANT: Banchereau, Jacques  
; TITLE OF INVENTION: PURIFIED MAMMALIAN CTLA-8 ANTIGENS AND  
; RELATED REAGENTS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DNAX Research Institute  
; STREET: 901 California Avenue  
; CITY: Palo Alto

STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/432,994A  
FILING DATE: 02-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/250,846  
FILING DATE: 27-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/177,747  
FILING DATE: 05-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,203  
FILING DATE: 14-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0388K3  
TELEPHONE: 415-852-9196  
TELEFAX: 415-496-1200  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 150 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-432-994A-2

Query Match 3.8%; Score 6; DB 4; Length 150;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
Db 137 VGCTCV 142

RESULT 10  
US-08-620-694A-8  
Sequence 8, Application US/08620694A  
Patent No. 5869286  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,694A  
FILING DATE: 21 MARCH 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Herpesvirus Saimiri  
STRAIN: ORF13  
US-08-620-694A-8

Query Match 3.8%; Score 6; DB 2; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
Db 138 VGCTCV 143

RESULT 11  
US-09-034-810-6  
Sequence 6, Application US/09034810  
Patent No. 6043344  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: Kelleher, Kerry  
APPLICANT: Carlin, McKeough  
APPLICANT: Goldman, Samuel  
APPLICANT: Pittman, Debra  
APPLICANT: Mi, Sha  
APPLICANT: Neben, Steven  
APPLICANT: Giannotti, JoAnn  
APPLICANT: Golden/Fleet, Margaret  
TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/034,810  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/685,239  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: G15262  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-034-810-6

Query Match 3.8%; Score 6; DB 3; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
Db 138 VGCTCV 143

RESULT 12  
US-09-022-255-8  
Sequence 8, Application US/09022255  
Patent No. 6072033  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,255  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/620,694  
FILING DATE: 21 MARCH 1996  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein

HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Herpesvirus Saimiri  
STRAIN: ORF13  
US-09-022-255-8

Query Match 3.8%; Score 6; DB 3; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
Db 138 VGCTCV 143

RESULT 13  
US-09-022-696-8  
Sequence 8, Application US/09022696  
Patent No. 6072037  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,696  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Herpesvirus Saimiri  
STRAIN: ORF13  
US-09-022-696-8

Query Match 3.8%; Score 6; DB 3; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
 Db 138 VGCTCV 143

RESULT 14  
 US-08-685-239-6  
 ; Sequence 6, Application US/08685239  
 ; Patent No. 6074849  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jacobs, Kenneth  
 ; APPLICANT: Kelleher, Kerry  
 ; APPLICANT: Carlin, McKeough  
 ; APPLICANT: Goldman, Samuel  
 ; APPLICANT: Pittman, Debra  
 ; APPLICANT: Mi, Sha  
 ; APPLICANT: Neben, Steven  
 ; APPLICANT: Giannotti, JoAnn  
 ; APPLICANT: Golden/Fleet, Margaret  
 ; TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genetics Institute, Inc.  
 ; STREET: 87 CambridgePark Drive  
 ; CITY: Cambridge  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02140  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/685,239  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brown, Scott A.  
 ; REGISTRATION NUMBER: 32,724  
 ; REFERENCE/DOCKET NUMBER: G15262  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 498-8224  
 ; TELEFAX: (617) 876-5851  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 151 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-685-239-6

Query Match 3.8%; Score 6; DB 3; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
 Db 138 VGCTCV 143

RESULT 15  
 US-09-022-253-8  
 ; Sequence 8, Application US/09022253  
 ; Patent No. 6096305  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yao, Zhengbin  
 ; APPLICANT: Spriggs, Melanie

; APPLICANT: Fanslow, William  
 ; TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Immunex Corporation  
 ; STREET: 51 University Street  
 ; CITY: Seattle  
 ; STATE: WA  
 ; COUNTRY: USA  
 ; ZIP: 98101  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: Apple Power Macintosh  
 ; OPERATING SYSTEM: Apple Operating System 7.5.5  
 ; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/022,253  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/620,694  
 ; FILING DATE: 21-MARCH-1996  
 ; APPLICATION NUMBER: USSN 08/538,765  
 ; FILING DATE: 7 AUGUST 1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: USSN 08/410,535  
 ; FILING DATE: 23 MARCH 1995  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Perkins, Patricia Anne  
 ; REGISTRATION NUMBER: 34,695  
 ; REFERENCE/DOCKET NUMBER: 2617-B  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206)587-0430  
 ; TELEFAX: (206)  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 151 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: No. 6096305 Relevant  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Herpesvirus Saimiri  
 ; STRAIN: ORF13  
 ; US-09-022-253-8

Query Match 3.8%; Score 6; DB 3; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
 Db 138 VGCTCV 143

RESULT 16  
 US-09-022-260-8  
 ; Sequence 8, Application US/09022260  
 ; Patent No. 6100235  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yao, Zhengbin  
 ; APPLICANT: Spriggs, Melanie  
 ; APPLICANT: Fanslow, William  
 ; TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Immunex Corporation  
 ; STREET: 51 University Street

;  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/022,260  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/620,694  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USN 08/410,535  
; FILING DATE: 23 MARCH 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia Anne  
; REGISTRATION NUMBER: 34,695  
; REFERENCE/DOCKET NUMBER: 2617-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206)587-0430  
; TELEFAX: (206)  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 151 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Herpesvirus Saimiri  
; STRAIN: ORF13  
; US-09-022-260-8

Query Match 3.8%; Score 6; DB 3; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.5e-02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 123 VGCTCV 128  
Db 138 VGCTCV 143  
|||||  
  
RESULT 17  
US-09-022-259-8  
; Sequence 8, Application US/09022259  
; Patent No. 6191104  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 6191104el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5

;  
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/022,259  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/620,694  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USN 08/410,535  
; FILING DATE: 23 MARCH 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia Anne  
; REGISTRATION NUMBER: 34,695  
; REFERENCE/DOCKET NUMBER: 2617-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206)587-0430  
; TELEFAX: (206)  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 151 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Herpesvirus Saimiri  
; STRAIN: ORF13  
; US-09-022-259-8

Query Match 3.8%; Score 6; DB 4; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.5e-02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
Db 138 VGCTCV 143  
|||||

RESULT 18  
US-09-022-257-8  
; Sequence 8, Application US/09022257  
; Patent No. 6197525  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 6197525el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/022,257  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/620,694  
; FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Herpesvirus Saimiri  
STRAIN: ORF13  
US-09-022-257-8

Query Match 3.8%; Score 6; DB 4; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
DB 138 VGCTCV 143

RESULT 19  
US-08-432-994A-4  
Sequence 4, Application US/08432994A  
Patent No. 6274711  
GENERAL INFORMATION:  
APPLICANT: Golstein, Pierre  
APPLICANT: Rouvier, Eric  
APPLICANT: Fossiez, Francois  
APPLICANT: Lebecque, Serge J.E.  
APPLICANT: Djossou, Odile  
APPLICANT: Banchereau, Jacques  
TITLE OF INVENTION: PURIFIED MAMMALIAN CTLA-8 ANTIGENS AND  
RELATED REAGENTS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/432,994A  
FILING DATE: 02-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/250,846  
FILING DATE: 27-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/177,747  
FILING DATE: 05-JAN-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/077,203  
FILING DATE: 14-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0388K3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-852-9196  
TELEFAX: 415-496-1200  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-432-994A-4

Query Match 3.8%; Score 6; DB 4; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
DB 138 VGCTCV 143

RESULT 20  
US-08-514-014-12  
Sequence 12, Application US/08514014  
Patent No. 5707829  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: McCoy, John  
APPLICANT: Kelleher, Kerry  
APPLICANT: Carlin, McKeough  
TITLE OF INVENTION: DNA SEQUENCES AND SECRETED PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc. -- Legal Affairs  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/514,014  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: GI6000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 153 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-514-014-12

Query Match 3.8%; Score 6; DB 1; Length 153;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
Db 140 VGCTCV 145

RESULT 21  
US-08-833-823-12  
; Sequence 12, Application US/08833823  
; Patent No. 5969093  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John  
; APPLICANT: Kelleher, Kerry  
; APPLICANT: Carlin, McKough  
; TITLE OF INVENTION: DNA SEQUENCES AND SECRETED PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc. -- Legal Affairs  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/833,823  
; FILING DATE: 10-APR-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/514,014  
; FILING DATE: 11-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brown, Scott A.  
; REGISTRATION NUMBER: 32,724  
; REFERENCE/DOCKET NUMBER: GI6000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8224  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 153 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-833-823-12

Query Match 3.8%; Score 6; DB 2; Length 153;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
Db 140 VGCTCV 145

RESULT 22  
US-08-432-944A-8  
; Sequence 8, Application US/08432994A  
; Patent No. 6274711  
; GENERAL INFORMATION:  
; APPLICANT: Golstein, Pierre  
; APPLICANT: Rouvier, Eric  
; APPLICANT: Fossiez, Francois  
; APPLICANT: Lebecque, Serge J.E.  
; APPLICANT: Djossou, Odile

; APPLICANT: Bancheureau, Jacques  
; TITLE OF INVENTION: PURIFIED MAMMALIAN CTLA-8 ANTIGENS AND  
; RELATED REAGENTS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DNAX Research Institute  
; STREET: 901 California Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/432,994A  
; FILING DATE: 02-MAY-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/250,846  
; FILING DATE: 27-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/177,747  
; FILING DATE: 05-JAN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/077,203  
; FILING DATE: 14-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ching, Edwin P.  
; REGISTRATION NUMBER: 34,090  
; REFERENCE/DOCKET NUMBER: DX0388K3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-852-9196  
; TELEFAX: 415-496-1200  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 155 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-432-994A-8

Query Match 3.8%; Score 6; DB 4; Length 155;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
Db 142 VGCTCV 147

RESULT 23  
US-08-620-694A-7  
; Sequence 7, Application US/08620694A  
; Patent No. 5869286  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,694A  
FILING DATE: 21 MARCH 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
CLASSIFICATION: 435  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-620-694A-7

Query Match 3.8%; Score 6; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 123 VGCTCV 128  
|||||  
Db 145 VGCTCV 150

RESULT 24  
US-09-022-255-7  
; Sequence 7, Application US/09022255  
; Patent No. 6072033  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,255  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/620,694  
FILING DATE: 21 MARCH 1996  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535

FILING DATE: 23 MARCH 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-255-7  
  
Query Match 3.8%; Score 6; DB 3; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 123 VGCTCV 128  
|||||  
Db 145 VGCTCV 150  
  
RESULT 25  
US-09-022-696-7  
; Sequence 7, Application US/09022696  
; Patent No. 6072037  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,696  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-09-022-696-7

Query Match 3.8%; Score 6; DB 3; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
Db 145 VGCTCV 150

RESULT 26

US-09-022-253-7  
; Sequence 7, Application US/09022253  
; Patent No. 6096305  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 6096305el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/022,253  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/620,694  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia Anne  
; REGISTRATION NUMBER: 34,695  
; REFERENCE/DOCKET NUMBER: 2617-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206)587-0430  
; TELEFAX: (206)  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 158 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-022-253-7

Query Match 3.8%; Score 6; DB 3; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
Db 145 VGCTCV 150

RESULT 27

US-09-022-260-7  
; Sequence 7, Application US/09022260  
; Patent No. 6100235  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 6100235el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Power Macintosh  
; OPERATING SYSTEM: Apple Operating System 7.5.5  
; SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/022,260  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/620,694  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 08/410,535  
; FILING DATE: 23 MARCH 1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia Anne  
; REGISTRATION NUMBER: 34,695  
; REFERENCE/DOCKET NUMBER: 2617-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206)587-0430  
; TELEFAX: (206)  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 158 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-022-260-7

Query Match 3.8%; Score 6; DB 3; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
|||||  
Db 145 VGCTCV 150

RESULT 28

US-09-022-259-7  
; Sequence 7, Application US/09022259  
; Patent No. 6191104  
; GENERAL INFORMATION:  
; APPLICANT: Yao, Zhengbin  
; APPLICANT: Spriggs, Melanie  
; APPLICANT: Fanslow, William  
; TITLE OF INVENTION: No. 6191104el Receptor That Binds IL-17  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle

STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,259  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-259-7

Query Match 3.8%; Score 6; DB 4; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
DB 145 VGCTCV 150

RESULT 29  
US-09-022-257-7  
Sequence 7, Application US/09022257  
Patent No. 6197525  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6197525el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,257  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-257-7

Query Match 3.8%; Score 6; DB 4; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
DB 145 VGCTCV 150

RESULT 30  
US-08-432-994A-10  
Sequence 10, Application US/08432994A  
Patent No. 6274711  
GENERAL INFORMATION:  
APPLICANT: Golstein, Pierre  
APPLICANT: Rouvier, Eric  
APPLICANT: Fossiez, Francois  
APPLICANT: Lebecque, Serge J.E.  
APPLICANT: Djossou, Odile  
APPLICANT: Banchereau, Jacques  
TITLE OF INVENTION: PURIFIED MAMMALIAN CTLA-8 ANTIGENS AND  
RELATED REAGENTS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/432,994A  
FILING DATE: 02-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/250,846  
FILING DATE: 27-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/177,747  
FILING DATE: 05-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/077,203  
FILING DATE: 14-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090

; REFERENCE/DOCKET NUMBER: DX0388K3  
; TELEPHONE: 415-852-9196  
; TELEFAX: 415-496-1200  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 158 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-432-994A-10

Query Match 3.8%; Score 6; DB 4; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
Db 145 VGCTCV 150

RESULT 31  
US-09-034-810-2  
; Sequence 2, Application US/09034810  
; Patent No. 6043344  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: Kelleher, Kerry  
; APPLICANT: Carlin, McKeough  
; APPLICANT: Goldman, Samuel  
; APPLICANT: Pittman, Debra  
; APPLICANT: Mi, Sha  
; APPLICANT: Neben, Steven  
; APPLICANT: Giannotti, JoAnn  
; APPLICANT: Golden/Fleet, Margaret  
; TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/034,810  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/685,239  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brown, Scott A.  
; REGISTRATION NUMBER: 32,724  
; REFERENCE/DOCKET NUMBER: G15262  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8224  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 163 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-034-810-2

Query Match 3.8%; Score 6; DB 3; Length 163;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
Db 150 VGCTCV 155

RESULT 32  
US-08-685-239-2  
; Sequence 2, Application US/08685239  
; Patent No. 6074849  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: Kelleher, Kerry  
; APPLICANT: Carlin, McKeough  
; APPLICANT: Goldman, Samuel  
; APPLICANT: Pittman, Debra  
; APPLICANT: Mi, Sha  
; APPLICANT: Neben, Steven  
; APPLICANT: Giannotti, JoAnn  
; APPLICANT: Golden/Fleet, Margaret  
; TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,239  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brown, Scott A.  
; REGISTRATION NUMBER: 32,724  
; REFERENCE/DOCKET NUMBER: G15262  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8224  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 163 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-685-239-2

Query Match 3.8%; Score 6; DB 3; Length 163;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 VGCTCV 128  
Db 150 VGCTCV 155

RESULT 33  
US-09-042-771-4  
; Sequence 4, Application US/09042771  
; Patent No. 6080729  
; GENERAL INFORMATION:  
; APPLICANT: Jaworski, Deborah  
; APPLICANT: Wang, Min

APPLICANT: Shilling, Lisa K.  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: spo0J2  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: US  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/042,771  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Falk, Stephen T  
REGISTRATION NUMBER: 36,795  
REFERENCE/DOCKET NUMBER: GM10135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-994-2488  
TELEFAX: 215-994-2222  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 179 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-042-771-4

Query Match 3.8%; Score 6; DB 3; Length 179;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ARAVLS 10  
| | | | |  
Db 86 ARAVLS 91

RESULT 34  
US-08-415-788-33  
Sequence 33, Application US/08415788  
Patent No. 5834591  
GENERAL INFORMATION:  
APPLICANT: NORMARK, STAFFAN  
APPLICANT: JONSSON, ANN-BETH  
TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES USEFUL  
FOR THE DIAGNOSIS AND TREATMENT OF PATHOGENIC NEISSERIA  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/415,788

FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/829,465  
FILING DATE: 31-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: SCHWENNING, LYNN E.  
REGISTRATION NUMBER: 37,233  
REFERENCE/DOCKET NUMBER: 29500-20046.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 244 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-415-788-33

Query Match 3.8%; Score 6; DB 2; Length 244;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 PTVVLR 102  
| | | | |  
Db 177 PTVVLR 182

RESULT 35  
US-08-415-788-31  
Sequence 31, Application US/08415788  
Patent No. 5834591  
GENERAL INFORMATION:  
APPLICANT: NORMARK, STAFFAN  
APPLICANT: JONSSON, ANN-BETH  
TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES USEFUL  
FOR THE DIAGNOSIS AND TREATMENT OF PATHOGENIC NEISSERIA  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/415,788  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/829,465  
FILING DATE: 31-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: SCHWENNING, LYNN E.  
REGISTRATION NUMBER: 37,233  
REFERENCE/DOCKET NUMBER: 29500-20046.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

; MOLECULE TYPE: protein  
; US-08-415-788-31

Query Match 3.8%; Score 6; DB 2; Length 246;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 97 PTVLVR 102  
|||||  
DB 179 PTVLVR 184

## RESULT 36

US-07-857-224B-17  
; Sequence 17, Application US/07857224B  
; Patent No. 5958784  
; GENERAL INFORMATION:  
; APPLICANT: Benner, Steven A.  
; TITLE OF INVENTION: Predicting Folded Structures of Proteins  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Steven A. Benner  
; STREET: Hadlaubstrasse 151  
; CITY: Zurich  
; STATE: none  
; COUNTRY: Switzerland  
; ZIP: (note: this is an international post code) CH-8092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh 7.0  
; SOFTWARE: Microsoft Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/857,224B  
; FILING DATE: 03/25/92  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA: none  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (International) 41 1 632 2830  
; TELEFAX: (International) 41 1 262 2437  
; TELEX: none  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 264  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Drosophila melanogaster  
; FEATURE: Protein kinase; Table 8 Column 18  
; PUBLICATION INFORMATION:  
; AUTHORS:  
; AUTHORS: Hanks, S. K.  
; AUTHORS: Quinn, A. M.  
; AUTHORS: Hunter, T.  
; TITLE: The protein kinase family  
; JOURNAL: Science  
; VOLUME: 241  
; PAGES: 42-52  
; DATE: 1988  
; US-07-857-224B-17

Query Match 3.8%; Score 6; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 GEEDVR 88  
|||||  
DB 251 GEEDVR 256

## RESULT 37

US-08-362-525-8  
; Sequence 8, Application US/08362525  
; Patent No. 6027910  
; GENERAL INFORMATION:  
; APPLICANT: KLIS, FRANCISCUS M.  
; APPLICANT: SCHREUDER, MAARTEN P.  
; APPLICANT: TOSCHKA, HOLSER V.  
; APPLICANT: VERRIPS, CORNELIS T.  
; TITLE OF INVENTION: PROCESS FOR IMMOBILIZING ENZYMES TO THE  
; TITLE OF INVENTION: CELL WALL OF A MICROBIAL CELL BY PRODUCING A FUSION  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,525  
; FILING DATE: 04-JAN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 92202080.5  
; FILING DATE: 08-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 92203899.7  
; FILING DATE: 14-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP93/01763  
; FILING DATE: 07-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOKULIS, PAUL N.  
; REGISTRATION NUMBER: 16,773  
; REFERENCE/DOCKET NUMBER: 213289/T7020(V)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 270 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-362-525-8

Query Match 3.8%; Score 6; DB 3; Length 270;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 154 NDAPAG 159  
|||||  
DB 27 NDAPAG 32

## RESULT 38

US-08-397-411-6  
; Sequence 6, Application US/08397411  
; Patent No. 6129914  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, George  
; APPLICANT: Gingrich, Roger  
; APPLICANT: Link, Brian

APPLICANT: Tso, J. Yun  
TITLE OF INVENTION: Bispecific Antibody Effective to Treat  
B-Cell Lymphoma and Cell Line  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/397,411  
FILING DATE: 01-MAR-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/859,583  
FILING DATE: 27-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 011823-004901  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 273 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-397-411-6

Query Match 3.8%; Score 6; DB 3; Length 273;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 CPAGGR 34  
Db 230 CPAGGR 235

RESULT 39  
US-09-042-771-2  
Sequence 2, Application US/09042771  
Patent No. 6080729  
GENERAL INFORMATION:  
APPLICANT: Jaworski, Deborah  
APPLICANT: Wang, Min  
APPLICANT: Shilling, Lisa K.  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: sp00j2  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: US  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/042,771  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Falk, Stephen T  
REGISTRATION NUMBER: 36,795  
REFERENCE/DOCKET NUMBER: GM10135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-994-2488  
TELEFAX: 215-994-2222  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 279 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-042-771-2

Query Match 3.8%; Score 6; DB 3; Length 279;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ARAVLS 10  
Db 186 ARAVLS 191

RESULT 40  
PCT-US95-04801-7  
Sequence 7, Application PC/TUS9504801  
GENERAL INFORMATION:  
APPLICANT: Martin, Juan F.  
APPLICANT: Coque, Juan R.  
APPLICANT: Enguita, Francisco J.  
APPLICANT: Fuente, Juan L.  
APPLICANT: Llaena, Francisco J.  
APPLICANT: Liras, Paloma  
TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John W. Wallen III  
STREET: P.O. Box 2000  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07065  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04801  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wallen III, John W.  
REGISTRATION NUMBER: 35,403  
REFERENCE/DOCKET NUMBER: 19179  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-3905  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 288 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-04801-7

Query Match 3.8%; Score 6; DB 5; Length 288;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SARARA 7  
Db 139 SARARA 144  
|||||

RESULT 41  
US-08-434-255-13  
; Sequence 13, Application US/08434255  
; Patent No. 5621089  
; GENERAL INFORMATION:  
; APPLICANT: Sloma, Alan P.  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Dambmann, Claus  
; APPLICANT: Aaslyng, Dorrit  
; TITLE OF INVENTION: ALKALINE PROTEASE  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5621089o No. 5621089th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Agis Dr., Cheryl H.  
; REGISTRATION NUMBER: 34,086  
; REFERENCE/DOCKET NUMBER: 3764.400-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-434-255-13

Query Match 3.8%; Score 6; DB 1; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
Db 48 NDAPAG 53  
|||||

RESULT 42  
US-08-459-967-13  
; Sequence 13, Application US/08459967  
; Patent No. 5622841  
; GENERAL INFORMATION:  
; APPLICANT: Sloma, Alan P.  
; APPLICANT: Outtrup, Helle

; APPLICANT: Dambmann, Claus  
; APPLICANT: Aaslyng, Dorrit  
; TITLE OF INVENTION: ALKALINE PROTEASE  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5622841o No. 5622841th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/434,255  
; FILING DATE: 03-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Agis Dr., Cheryl H.  
; REGISTRATION NUMBER: 34,086  
; REFERENCE/DOCKET NUMBER: 3764.400-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-459-967-13

Query Match 3.8%; Score 6; DB 1; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
Db 48 NDAPAG 53  
|||||

RESULT 43  
US-08-460-327-13  
; Sequence 13, Application US/08460327  
; Patent No. 5622850  
; GENERAL INFORMATION:  
; APPLICANT: Sloma, Alan P.  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Dambmann, Claus  
; APPLICANT: Aaslyng, Dorrit  
; TITLE OF INVENTION: ALKALINE PROTEASE  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5622850o No. 5622850th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,327

;; FILING DATE: 02-JUN-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/434,255  
;; FILING DATE: 03-MAY-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Agris Dr., Cheryl H.  
;; REGISTRATION NUMBER: 34,086  
;; REFERENCE/DOCKET NUMBER: 3764.400-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-867-0123  
;; TELEFAX: 212-878-9655  
;; INFORMATION FOR SEQ ID NO: 13:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 291 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-460-327-13

Query Match 3.8%; Score 6; DB 1; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
Db 48 NDAPAG 53

## RESULT 44

US-08-459-871-13  
; Sequence 13, Application US/08459871  
; Patent No. 5650326  
; GENERAL INFORMATION:  
; APPLICANT: Sloma, Alan P.  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Dambmann, Claus  
; APPLICANT: Aaslyng, Dorrit  
; TITLE OF INVENTION: ALKALINE PROTEASE  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 56503260 No. 5650326disk of No. 5650326th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,871  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/434,255  
; FILING DATE: 03-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Agris Dr., Cheryl H.  
; REGISTRATION NUMBER: 34,086  
; REFERENCE/DOCKET NUMBER: 3764.400-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-459-871-13

Query Match 3.8%; Score 6; DB 1; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
Db 48 NDAPAG 53

## RESULT 45

US-08-244-686-4  
; Sequence 4, Application US/08244686  
; Patent No. 5705358  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Process for the production of a protein  
; using endoxylanase II (exIA) expression signals  
; NUMBER OF SEQUENCES: 12  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/244,686  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-244-686-4

Query Match 3.8%; Score 6; DB 1; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
Db 48 NDAPAG 53

## RESULT 46

US-08-784-651-16  
; Sequence 16, Application US/08784651  
; Patent No. 5821102  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy  
; APPLICANT: Boominathan, Karuppan  
; APPLICANT: Sandal, Thomas  
; TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5821102o No. 5821102disk of No. 5821102th America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/784,651  
; FILING DATE: 21-JAN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:

; NAME: Agris, Cheryl H.  
; REGISTRATION NUMBER: 34,086  
; REFERENCE/DOCKET NUMBER: 4608.200-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-784-651-16

Query Match 3.8%; Score 6; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
Db 48 NDAPAG 53  
|||||

RESULT 47  
US-08-921-426-10  
; Sequence 10, Application US/08921426  
; Patent No. 5837847  
; GENERAL INFORMATION:  
; APPLICANT: Royer, John C  
; APPLICANT: Moyer, Donna L  
; APPLICANT: Yoder, Wendy T  
; APPLICANT: Shuster, Jeffrey R  
; TITLE OF INVENTION: NON-TOXIC, NON-TOXIGENIC, NON-PATHOGENIC  
; TITLE OF INVENTION: FUSARIUM EXPRESSION SYSTEM  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5837847o No. 5837847disk of No. 5837847th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/921,426  
; FILING DATE: 29-AUG-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/456,433  
; FILING DATE: 01-JUN-1995  
; APPLICATION NUMBER: US 08/404,678  
; FILING DATE: 15-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Agris Dr., Cheryl H.  
; REGISTRATION NUMBER: 34,086  
; REFERENCE/DOCKET NUMBER: 4216.010-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-921-426-10

Query Match 3.8%; Score 6; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
Db 48 NDAPAG 53  
|||||

RESULT 48  
US-08-479-275D-2  
; Sequence 2, Application US/08479275D  
; Patent No. 5869438  
; GENERAL INFORMATION:  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Patkar, Shamkant A.  
; APPLICANT: Gormsen, Erik  
; APPLICANT: Clausen, Ib G.  
; APPLICANT: Okkels, Jens S.  
; APPLICANT: Thellersen, Marianne  
; TITLE OF INVENTION: LIPASE VARIANTS  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5869438o No. 5869438disk of No. 5869438th America, Inc.  
; STREET: 405 Lexington Avenue, Suite 6400  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/479,275D  
; FILING DATE: 07-JUNE-1995  
; CLASSIFICATION: 510  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 3520.514-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-479-275D-2

Query Match 3.8%; Score 6; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
Db 48 NDAPAG 53  
|||||

RESULT 49  
US-08-488-271B-2  
; Sequence 2, Application US/08488271B  
; Patent No. 5892013  
; GENERAL INFORMATION:  
; APPLICANT: Svendsen, Allan  
; APPLICANT: Patkar, Shamkant A.  
; APPLICANT: Gormsen, Erik  
; APPLICANT: Clausen, Ib G.  
; APPLICANT: Okkels, Jens S.

; APPLICANT: Thellersen, Marianne  
; TITLE OF INVENTION: LIPASE VARIANTS  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 58920130 No. 5892013disk of No. 5892013th America, Inc.  
; STREET: 405 Lexington Avenue, Suite 6400  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174-6201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,271B  
; FILING DATE: 07-JUNE-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 3520.504-US  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-488-271B-2

Query Match 3.8%; Score 6; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
| | | | |  
Db 48 NDAPAG 53

RESULT 50  
US-08-140-008A-2  
; Sequence 2, Application US/08140008A  
; Patent No. 5914306  
; GENERAL INFORMATION:  
; APPLICANT: SVENDSEN, Allan  
; APPLICANT: VON DER OSTEN, Claus  
; APPLICANT: CLAUSEN, Ib Groth  
; APPLICANT: PATKAR, Shamkant Anant  
; APPLICANT: BORCH, Kim  
; TITLE OF INVENTION: STABILIZED ENZYMES  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 59143060 No. 5914306disk of No. 5914306th America  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/140,008A  
; FILING DATE: 22-OCT-1993  
; CLASSIFICATION: 252  
; ATTORNEY/AGENT INFORMATION:

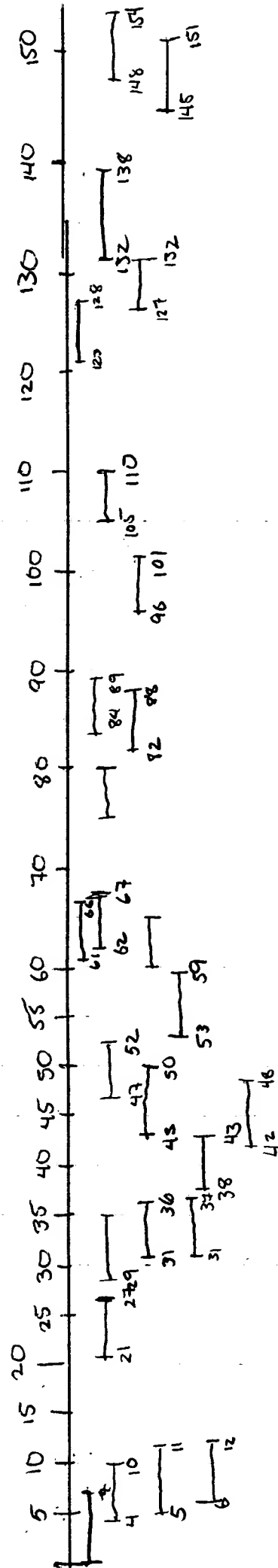
; NAME: Lambiris, Elias J.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 3601.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 867-0123  
; TELEFAX: (212) 878-9655  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-140-008A-2

Query Match 3.8%; Score 6; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 NDAPAG 159  
| | | | |  
Db 48 NDAPAG 53

Search completed: April 14, 2002, 12:42:29  
Job time: 201 sec

---



**THIS PAGE BLANK (USPTO)**



8b/1/t  
9b/1/11  
8b/1/t  
9b/1/11  
9b/1/11  
tb/1/s  
tb/1/s  
1/1/98

990470	98-86
681698	97-103
1043701	26-32
04740	145-151
900150	21-27
96356	42-48
84378	31-37
029550	6-12

09/731816-3

**THIS PAGE BLANK (USPTO)**